



# Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling

Shufen Pan<sup>1</sup>, Naiqing Pan<sup>1,2</sup>, Hanqin Tian<sup>1</sup>, Pierre Friedlingstein<sup>3</sup>, Stephen Sitch<sup>4</sup>, Hao Shi<sup>1</sup>, Vivek K. Arora<sup>5</sup>, Vanessa Haverd<sup>6</sup>, Atul K. Jain<sup>7</sup>, Etsushi Kato<sup>8</sup>, Sebastian Lienert<sup>9</sup>, Danica Lombardozzi<sup>10</sup>, Julia E. M. S. Nabel<sup>11</sup>, Catherine Ottlé<sup>12</sup>, Benjamin Poulter<sup>13</sup>, Sönke Zaehle<sup>14</sup>, and Steven W. Running<sup>15</sup>

<sup>1</sup>International Center for Climate and Global Change Research, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL 36832, USA

<sup>2</sup>State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

<sup>3</sup>College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK

<sup>4</sup>College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4RJ, UK

<sup>5</sup>Canadian Centre for Climate Modelling and Analysis, Environment Canada, University of Victoria, Victoria, BC V8W 2Y2, Canada

<sup>6</sup>CSIRO Oceans and Atmosphere, GPO Box 1700, Canberra, ACT 2601, Australia

<sup>7</sup>Department of Atmospheric Sciences, University of Illinois, Urbana, IL 61801, USA

<sup>8</sup>Institute of Applied Energy (IAE), Minato-ku, Tokyo 105-0003, Japan

<sup>9</sup>Climate and Environmental Physics, Physics Institute, University of Bern, Bern, Switzerland

<sup>10</sup>Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO 80305, USA

<sup>11</sup>Max Planck Institute for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany

<sup>12</sup>LSCE-IPSL-CNRS, Orme des Merisiers, 91191, Gif-sur-Yvette, France

<sup>13</sup>NASA Goddard Space Flight Center, Biospheric Sciences Laboratory, Greenbelt, MD 20771, USA

<sup>14</sup>Max Planck Institute for Biogeochemistry, P.O. Box 600164, Hans-Knöll-Str. 10, 07745 Jena, Germany

<sup>15</sup>Numerical Terradynamic Simulation Group, College of Forestry and Conservation, University of Montana, Missoula, MT 59812, USA

**Correspondence:** Shufen Pan (panshuf@auburn.edu)

Received: 3 August 2019 – Discussion started: 21 August 2019

Revised: 22 January 2020 – Accepted: 27 January 2020 – Published: 31 March 2020

**Abstract.** Evapotranspiration (ET) is critical in linking global water, carbon and energy cycles. However, direct measurement of global terrestrial ET is not feasible. Here, we first reviewed the basic theory and state-of-the-art approaches for estimating global terrestrial ET, including remote-sensing-based physical models, machine-learning algorithms and land surface models (LSMs). We then utilized 4 remote-sensing-based physical models, 2 machine-learning algorithms and 14 LSMs to analyze the spatial and temporal variations in global terrestrial ET. The results showed that the ensemble means of annual global terrestrial ET estimated by

these three categories of approaches agreed well, with values ranging from  $589.6 \text{ mm yr}^{-1}$  ( $6.56 \times 10^4 \text{ km}^3 \text{ yr}^{-1}$ ) to  $617.1 \text{ mm yr}^{-1}$  ( $6.87 \times 10^4 \text{ km}^3 \text{ yr}^{-1}$ ). For the period from 1982 to 2011, both the ensembles of remote-sensing-based physical models and machine-learning algorithms suggested increasing trends in global terrestrial ET ( $0.62 \text{ mm yr}^{-2}$  with a significance level of  $p < 0.05$  and  $0.38 \text{ mm yr}^{-2}$  with a significance level of  $p < 0.05$ , respectively). In contrast, the ensemble mean of the LSMs showed no statistically significant change ( $0.23 \text{ mm yr}^{-2}$ ,  $p > 0.05$ ), although many of the individual LSMs reproduced an increasing trend. Nevertheless,

all 20 models used in this study showed that anthropogenic Earth greening had a positive role in increasing terrestrial ET. The concurrent small interannual variability, i.e., relative stability, found in all estimates of global terrestrial ET, suggests that a potential planetary boundary exists in regulating global terrestrial ET, with the value of this boundary being around  $600 \text{ mm yr}^{-1}$ . Uncertainties among approaches were identified in specific regions, particularly in the Amazon Basin and arid/semiarid regions. Improvements in parameterizing water stress and canopy dynamics, the utilization of new available satellite retrievals and deep-learning methods, and model–data fusion will advance our predictive understanding of global terrestrial ET.

## 1 Introduction

Terrestrial evapotranspiration (ET) is the sum of the water lost to the atmosphere from plant tissues via transpiration and that lost from the land surface elements, including soil, plants and open water bodies, through evaporation. Processes controlling ET play a central role in linking the energy (latent heat), water (moisture flux) and carbon cycles (photosynthesis–transpiration trade-off) in the Earth system. Over 60 % of precipitation on the land surface is returned to the atmosphere through ET (Oki and Kanae, 2006), and the accompanying latent heat ( $\lambda \text{ET}$ ,  $\lambda$  is the latent heat of vaporization) accounts for more than half of the solar energy received by the land surface (Trenberth et al., 2009). ET is also coupled with the carbon dioxide ( $\text{CO}_2$ ) exchange between the canopy and the atmosphere through vegetation photosynthesis. These linkages make ET an important variable in both short-term numerical weather forecasts and long-term climate predictions. Moreover, ET is a critical indicator for ecosystem functioning across a variety of spatial scales. Therefore, in order to enhance our predictive understanding of the Earth system and sustainability, it is essential to accurately assess land surface ET in a changing global environment.

However, large uncertainty still exists in quantifying the magnitude of global terrestrial ET and its spatial and temporal patterns, despite extensive research (Allen et al., 1998; Liu et al., 2008; Miralles et al., 2016; Mueller et al., 2011; Tian et al., 2010). Previous estimates of global land mean annual ET range from 417 to  $650 \text{ mm yr}^{-1}$  for the whole or part of the 1982–2011 period (Mu et al., 2007; Mueller et al., 2011; Vinukollu et al., 2011a; Zhang et al., 2010). This large discrepancy among independent studies may be attributed to a lack of sufficient measurements, uncertainty in forcing data, inconsistent spatial and temporal resolutions, ill-calibrated model parameters, and deficiencies in model structures. Of the four components of ET (transpiration, soil evaporation, canopy interception and open water evaporation), transpiration ( $T_v$ ) contributes the largest uncertainty;

this is due to the fact that it is modulated not only by surface meteorological conditions and soil moisture but also by the physiology and structures of plants. Changes in non-climatic factors such as elevated atmospheric  $\text{CO}_2$ , nitrogen deposition and land covers also serve as influential drivers of  $T_v$  (Gedney et al., 2006; Mao et al., 2015; S. Pan et al., 2018; Piao et al., 2010). As such, the global ratio of transpiration to ET ( $T_v/\text{ET}$ ) has long been a matter of debate, with the most recent observation-based estimate being  $0.64 \pm 0.13$  constrained by the global water-isotope budget (Good et al., 2015). Most Earth system models are thought to largely underestimate  $T_v/\text{ET}$  (Lian et al., 2018).

Global warming is expected to accelerate the hydrological cycle (Pan et al., 2015). For the period from 1982 to the late 1990s, ET was reported to have increased by about 7 mm ( $\sim 1.2\%$ ) per decade driven by an increase in radiative forcing and, consequently, global and regional temperatures (Douville et al., 2013; Jung et al., 2010; Wang et al., 2010). The contemporary near-surface specific humidity also increased over both land and ocean (Dai, 2006; Simmons et al., 2010; Willett et al., 2007). More recent studies confirmed that global ET has showed an overall increase since the 1980s (Mao et al., 2015; Yao et al., 2016; Zeng et al., 2018a, 2012, 2016; Zhang et al., 2015; Y. Zhang et al., 2016). However, the magnitude and spatial distribution of such a trend are far from determined. Over the past 50 years, pan evaporation has decreased worldwide (Fu et al., 2009; Peterson et al., 1995; Roderick and Farquhar, 2002), implying an increase in actual ET given the pan evaporation paradox. Moreover, the increase in global terrestrial ET was found to cease or even be reversed from 1998 to 2008, primarily due to the decreased soil moisture supply in the Southern Hemisphere (Jung et al., 2010). To reconcile the disparity, Douville et al. (2013) argued that the peak ET in 1998 should not be taken as a tipping point because ET was estimated to increase in the multi-decadal evolution. More efforts are needed to understand the spatial and temporal variations of global terrestrial ET and the underlying mechanisms that control its magnitude and variability.

Conventional techniques, such as lysimeter, eddy covariance, large aperture scintillometer and the Bowen ratio method, are capable of providing ET measurements at point and local scales (Wang and Dickinson, 2012). However, it is impossible to directly measure ET at the global scale because dense global coverage by such instruments is not feasible, and the representativeness of point-scale measurements to comprehensively portray the spatial heterogeneity of global land surface is also doubtful (Mueller et al., 2011). To address this issue, numerous approaches have been proposed in recent years to estimate global terrestrial ET and these approaches can be divided into three main categories: (1) remote-sensing-based physical models, (2) machine-learning algorithms and (3) land surface models (Miralles et al., 2011; Mueller et al., 2011; Wang and Dickinson, 2012). Knowledge of the uncertainties in global ter-

restrial ET estimates from different approaches is a prerequisite for future projection and many other applications. In recent years, several studies have compared multiple terrestrial ET estimates (Khan et al., 2018; Mueller et al., 2013; Wartenburger et al., 2018; Y. Zhang et al., 2016); however, most of these studies analyzed multiple datasets of the same approach or focused on investigating similarities and differences among different approaches. Few studies have been conducted to identify uncertainties in multiple estimates of different approaches.

In this study, we integrate state-of-the-art estimates of global terrestrial ET, including data-driven and process-based estimates, to assess its spatial pattern, interannual variability, environmental drivers, long-term trend and response to vegetation greening. Our goal is not to compare the various models and choose the best one but to identify the uncertainty sources in each type of estimate and provide suggestions for future model development. In the following sections, we first give a brief introduction to all of the methodological approaches and ET datasets used in this study. We then quantify the spatiotemporal variations in global terrestrial ET during the period from 1982 to 2011 by analyzing the results from the current state-of-the-art models. Finally, we discuss some suggested solutions for reducing the identified uncertainties.

## 2 Methodology and data sources

### 2.1 Overview of approaches to global ET estimation

#### 2.1.1 Remote-sensing-based physical models

Satellite remote sensing has been widely recognized as a promising tool for estimating global ET, because it is capable of providing spatially and temporally continuous measurements of critical biophysical parameters affecting ET, including vegetation states, albedo, the fraction of absorbed photosynthetically active radiation, land surface temperature and plant functional types (Li et al., 2009). Since the 1980s, a large number of methods have been developed using a variety of satellite observations (K. Zhang et al., 2016). However, some of these methods, such as surface energy balance (SEB) models and surface temperature–vegetation index (Ts–VI), are usually applied at local and regional scales. At the global scale, the vast majority of the existing remote-sensing-based physical models can be categorized into two groups: those based on the Penman–Monteith (PM) equation and those based on the Priestley–Taylor (PT) equation.

#### Remote sensing models based on the Penman–Monteith equation

The Penman equation, derived from the Monin–Obukhov similarity theory and surface energy balance, uses surface net radiation, temperature, humidity, wind speed and ground heat

flux to estimate ET from an open water surface. For vegetated surfaces, canopy resistance was introduced into the Penman equation by Monteith (Monteith, 1965), and the PM equation is formulated as follows:

$$\lambda ET = \frac{\Delta (R_n - G) + \rho_a C_p VPD / r_a}{\Delta + \gamma (1 + r_s / r_a)}, \quad (1)$$

where  $\Delta$ ,  $R_n$ ,  $G$ ,  $\rho_a$ ,  $C_p$ ,  $\gamma$ ,  $r_s$ ,  $r_a$  and VPD are the slope of the curve relating saturated water vapor pressure to air temperature, net radiation, soil heat flux, air density, the specific heat of air, the psychrometric constant, surface resistance, aerodynamic resistance and the vapor pressure deficit, respectively. The canopy resistance term in the PM equation exerts a strong control on transpiration. For example, based on the algorithm proposed by Cleugh et al. (2007), the MODIS (Moderate Resolution Imaging Spectroradiometer) ET algorithm improved the model performance via the inclusion of environmental stress into the canopy conductance calculation and explicitly accounted for soil evaporation (Mu et al., 2007). Further, Mu et al. (2011) improved the MODIS ET algorithm by considering nighttime ET, adding the soil heat flux calculation, separating the dry canopy surface from the wet, and dividing the soil surface into saturated wet surface and moist surface. Similarly, Zhang et al. (2010) developed a Jarvis–Stewart-type canopy conductance model based on the normalized difference vegetation index (NDVI) to take advantage of the long-term Advanced Very High Resolution Radiometer (AVHRR) dataset. More recently, this model was improved by adding a CO<sub>2</sub> constraint function into the canopy conductance estimate (Zhang et al., 2015). Another important revision for the PM approach is proposed by Leuning et al. (2008). The Penman–Monteith–Leuning method adopts a simple biophysical model for canopy conductance, which can account for the influences of radiation and the atmospheric humidity deficit. Additionally, it introduces a simpler soil evaporation algorithm than that proposed by Mu et al. (2007), which potentially makes it attractive for use with remote sensing. However, PM-based models have one intrinsic weakness – temporal upscaling – which is required when translating instantaneous ET estimation into a longer timescale value (Li et al., 2009). This could easily be done at the daily scale under clear-sky conditions but faces challenge at weekly to monthly timescales due to lack of cloud coverage information.

#### Remote sensing models based on the Priestley–Taylor equation

The Priestley–Taylor (PT) equation is a simplification of the PM equation that does not parameterize aerodynamic and surface conductance (Priestley and Taylor, 1972); it can be expressed as follows:

$$\lambda ET = f_{\text{stress}} \times \alpha \times \frac{\Delta}{\Delta + \gamma} \times (R_n - G), \quad (2)$$

where  $f_{\text{stress}}$  is a stress factor and is usually computed as a function of environmental conditions.  $\alpha$  is the PT parameter with a value of between 1.2 and 1.3 under water unstressed conditions and can be estimated using remote sensing. Although the original PT equation works well for estimating potential ET across most surfaces, the Priestley–Taylor coefficient,  $\alpha$ , usually needs adjustment to convert potential ET to actual ET (K. Zhang et al., 2016). Thus, Fisher et al. (2008) developed a modified PT model that keeps  $\alpha$  constant but scales down potential ET using eco-physiological constraints and soil evaporation partitioning. The accuracy of their model has been validated against eddy-covariance measurements conducted in a wide range of climates and involving many plant functional types (Fisher et al., 2009; Vinukollu et al., 2011b). Following this idea, Yao et al. (2013) further developed a modified Priestley–Taylor algorithm that constrains soil evaporation using the index of the soil water deficit derived from the apparent thermal inertia. Miralles et al. (2011) also proposed a novel PT-type model: the Global Land Evaporation Amsterdam Model (GLEAM). GLEAM combines a soil water module, a canopy interception model and a stress module within the PT equation. The key distinguishing features of this model are the use of microwave-derived soil moisture, land surface temperature and vegetation density, and the detailed estimation of rainfall interception loss. In this way, GLEAM minimizes the dependence on static variables, avoids the need for parameter tuning and enables the quality of the evaporation estimates to rely on the accuracy of the satellite inputs (Miralles et al., 2011). Compared with the PM approach, the PT-based approaches avoid the computational complexities of aerodynamic resistance and the accompanying error propagation. However, the many simplifications and semiempirical parameterization of physical processes in the PT-based approaches may lower its accuracy.

### 2.1.2 Vegetation-index-based empirical algorithms and machine-learning methods

The principle of empirical ET algorithms is to link observed ET to its controlling environmental factors through various statistical regressions or machine-learning algorithms of different complexities. The earliest empirical regression method was proposed by Jackson et al. (1977). At present, the majority of regression models are based on vegetation indices (Glenn et al., 2010), such as the NDVI and the enhanced vegetation index (EVI), because of their simplicity, resilience in the presence of data gaps, utility under a wide range of conditions and connection with vegetation transpiration capacity (Maselli et al., 2014; Nagler et al., 2005; Yuan et al., 2010). As an alternative to statistical regression methods, machine-learning algorithms have been gaining increased attention for ET estimation due to their ability to capture the complex non-linear relationships between ET and its controlling factors (Dou and Yang, 2018). Many conventional machine-learning

algorithms, such as those based on artificial neural networks, random forest, and support vector machine algorithms, have been applied in various ecosystems (Antonopoulos et al., 2016; Chen et al., 2014; Feng et al., 2017; Shrestha and Shukla, 2015) and have proved to be more accurate in estimating ET than simple regression models (Antonopoulos et al., 2016; Chen et al., 2014; Kisi et al., 2015; Shrestha and Shukla, 2015; Tabari et al., 2013). In upscaling FLUXNET ET to the global scale, Jung et al. (2010) used the model tree ensemble method to integrate eddy-covariance measurements of ET with satellite remote sensing and surface meteorological data. In a recent study (Bodesheim et al., 2018), the random forest approach was used to derive global ET at a 30 min timescale.

### 2.1.3 Process-based land surface models (LSMs)

Although satellite-derived ET products have provided quantitative investigations of historical terrestrial ET dynamics, they can only cover a limited temporal record of about 4 decades. To obtain terrestrial ET before the 1980s and predict future ET dynamics, LSMs are needed, as they are able to represent a large number of interactions and feedbacks between physical, biological and biogeochemical processes in a prognostic way (Jimenez et al., 2011). ET simulation in LSMs is regulated by multiple biophysical and physiological properties or processes, including but not limited to stomatal conductance, leaf area, root water uptake, soil water, runoff and (sometimes) nutrient uptake (Famiglietti and Wood, 1991; Huang et al., 2016; Lawrence et al., 2007). Although almost all current LSMs have these components, different parameterization schemes result in substantial differences in ET estimation (Wartenburger et al., 2018). Therefore, in recent years, the multi-model ensemble approach has become popular in quantifying the magnitude, spatiotemporal pattern and uncertainty of global terrestrial ET (Mueller et al., 2011; Wartenburger et al., 2018). Yao et al. (2017) showed that a simple model averaging method or a Bayesian model averaging method is superior to each individual model in predicting terrestrial ET.

## 2.2 Description of ET models used in this study

In this study, we evaluate 20 ET products that are based on remote-sensing-based physical models, machine-learning algorithms and LSMs to investigate the magnitudes and spatial patterns of global terrestrial ET over recent decades. Table 1 lists the input data, the ET algorithms adopted, the advantages and limitations, and the references for each product. We use a simple model averaging method when calculating the mean value of multiple models.

Four physically based remote sensing datasets, including the Process-based Land Surface Evapotranspiration/Heat Fluxes algorithm (P-LSH), the Global Land Evaporation Amsterdam Model (GLEAM), the Moderate Resolution

**Table 1.** Descriptions of the models used in this study, including their drivers, adopted algorithms, key equations, advantages and limitations, and references.

Name	Input	Algorithm	Spatial resolution	Temporal resolution	Key equations	Advantages and limitations	References
MTE	Climate: precipitation, temperature, sunshine hour, relative humidity and wet days Vegetation: $f_{\text{APAR}}$	TRIAL + ERROR	$0.5^\circ \times 0.5^\circ$	Monthly	No specific equation	Limitations: insufficient flux observations in tropical regions; no $\text{CO}_2$ effect	Jung et al. (2011)
RF	Climate: land surface temperature, radiation, potential radiation, index of water availability and relative humidity Vegetation: $f_{\text{APAR}}$ , LAI and enhanced vegetation index (EVI)	Randomized decision tree	$0.5^\circ \times 0.5^\circ$	30 min	No specific equation	Limitations: the same as for MTE	Bodesheim et al. (2018)
P-LSH	Climate: radiation, air temperature, vapor pressure, wind speed and $\text{CO}_2$ Vegetation: AVHRR NDVI	Modified Penman–Monteith	$0.083^\circ \times 0.083^\circ$	Monthly	$E_v = \frac{\Delta R_n + \rho C_p \text{VPD} g_a}{\lambda_v \left( \Delta + \gamma \left( 1 + \frac{g_a}{g_s} \right) \right)}$ $E_s = \text{RH} \frac{\text{VPD}}{k} \frac{\Delta R_n + \rho C_p \text{VPD} g_a}{\lambda_v \left( \Delta + \gamma \left( 1 + \frac{g_a}{g_s} \right) \right)}$	Advantages: more robust physical basis; considers the effects of $\text{CO}_2$ Limitations: high meteorological forcing requirements; canopy conductance is based on proxies	Zhang et al. (2015)
GLEAM	Climate: precipitation, net radiation, surface soil moisture, land surface temperature, air temperature and snow depth Vegetation: vegetation optical depth	Modified Priestley–Taylor	$0.25^\circ \times 0.25^\circ$	Daily	$E_s = f_s S \alpha_s \frac{\Delta}{\lambda_v \rho_w (\Delta + \gamma)} (R_n^s - G_s)$ $E_{sc} = f_{sc} S_{sc} \alpha_{sc} \frac{\Delta}{\lambda_v \rho_w (\Delta + \gamma)} (R_n^{sc} - G_{sc})$ $E_{tc} = f_{tc} S_{tc} \alpha_{tc} \frac{\Delta}{\lambda_v \rho_w (\Delta + \gamma)} (R_n^{tc} - G_{tc}) - \beta E_i$	Advantages: simple; low requirement for meteorological data; well-suited for remote sensing observable variables; soil moisture is considered Limitations: many simplifications of physical processes; neither VPD nor surface and aerodynamic resistances are explicitly accounted for; strong dependency on net radiation	Miralles et al. (2011)
MODIS	Climate: air temperature, shortwave radiation, wind speed, relative humidity and air pressure Vegetation: LAI and $f_{\text{APAR}}$ , albedo	Penman–Monteith–Leuning	$0.05^\circ \times 0.05^\circ$	Monthly	$E_i = f_{\text{wet}} f_c \frac{\Delta (R_n - G) + \rho C_p \frac{\text{VPD}}{r_a}}{\lambda_v \rho_w (\Delta + \gamma \frac{r_a}{r_{\text{wet}}})}$ $E_v = (1 - f_{\text{wet}}) f_c \frac{\Delta (R_n - G) + \rho C_p \frac{\text{VPD}}{r_a}}{\lambda_v \rho_w (\Delta + \gamma \frac{r_a}{r_a})}$ $E_s = \left[ f_{\text{wet}} + \frac{(1 - f_{\text{wet}}) h \text{VPD}}{\beta} \right] \frac{(s A_{\text{soil}} + \rho C_p (1 - f_c) \text{VPD})}{\lambda_v \rho_w (S + \gamma \frac{r_{\text{soil}}}{r_{\text{as}}})}$	Advantages: more robust physical basis; Limitations: requires many variables that are difficult to observe or that are not observable with satellites; canopy conductance is based on proxies; does not consider soil moisture but uses atmospheric humidity as a surrogate; does not consider the effects of $\text{CO}_2$	Mu et al. (2011)
PML-CSIRO	Climate: precipitation, air temperature, vapor pressure, shortwave radiation, longwave radiation and wind speed Vegetation: AVHRR LAI, emissivity and albedo	Penman–Monteith–Leuning	$0.5^\circ \times 0.5^\circ$	Monthly	$E_v = \frac{\Delta R_n + \rho C_p \text{VPD} g_a}{\lambda_v (\Delta + \gamma (1 + \frac{g_a}{g_s}))}$ $E_s = \frac{f \Delta A_s}{\Delta + \gamma}$ $E_i: \text{an adapted version of the Gash rainfall interception model (Van Dijk and Bruijnzeel, 2001)}$	Advantages: more robust physical basis (compared with Priestley–Taylor equation); biophysically based estimation of surface conductance Limitations: high meteorological forcing requirements; canopy conductance is based on proxies; does not consider the effects of $\text{CO}_2$	Y. Zhang et al. (2016)

Table 1. Continued.

TRENDY LSMs				
<p>Advantages: land surface models are process-oriented and physically based. Given their structure almost all models are capable of allowing factorial analysis, where one forcing can be applied at a time. Most models also consider the physiological effect of CO<sub>2</sub> on stomatal closure.</p> <p>Limitations: most models typically do not allow integration/assimilation of observation-based vegetation characteristics. Model parameterizations remain uncertain and the same process is modeled in different ways across models. Model parameters may or may not be physically based and, therefore, measurable in the field.</p> <p>Models participating in the TRENDY6 comparison were forced by precipitation, air temperature, specific humidity, shortwave radiation, longwave radiation and wind speed based on the CRU-NCAP8 data, as explained in Le Quéré et al. (2018). It is very difficult to list all key equations for all land surface models. Thus, we only list the stomatal conductance equation for each model in this table.</p>				
Name	Algorithm	Spatial resolution	Temporal resolution	Key equations
CABLE	Penman–Monteith	0.5° × 0.5°	Monthly	$g_s = g_0 + \frac{g_1 f_{sc} A}{c_a - c_p} \left(1 + \frac{VPD}{VPD_0}\right)^{-1}$
CLASS-CTEM	Modified Penman–Monteith	2.8125° × 2.8125°	Monthly	$g_s = m \frac{A_{LAI}}{(c_s - 1)} \frac{1}{(1 + VPD/VPD_0)} + b LAI$
CLM45	Modified Penman–Monteith	1.875° × 2.5°	Monthly	$g_s = g_0 + \frac{g_1 A}{c_a} RH$
DLEM	Penman–Monteith	0.5° × 0.5°	Monthly	$g_s = \max(g_{smax}, f_{corb} f(T_{min}) f(VPD) f(CO_2) \cdot g_{smin})$
ISAM	Modified Penman–Monteith	0.5° × 0.5°	Monthly	$g_s = m \frac{A}{c_s / P_{sat}} \times \frac{c_s}{c_i} + b \beta_i$
JSBACH	Penman–Monteith	1.9° × 1.9°	Monthly	$g_s = \beta_w \frac{1.6 A_{LAI}}{c_a - c_{i,pot}}$
JULES	Penman–Monteith	2.5° × 3.75°	Monthly	<p>Bare soil conductance: <math>g_{soil} = \frac{1}{100} \left(\frac{\theta_i}{\theta_c}\right)^2</math></p> <p>Stomatal conductance is calculated by solving the following two equations:</p> $A_l = g_s (C_s - C_l) / 1.6; \frac{C_i - \Gamma^*}{c_a - \Gamma^*} = f_0 \left(1 - \frac{\Delta}{q_c}\right)$
LPJ-GUESS	Equations proposed by Monteith (1995)	0.5° × 0.5°	Monthly	$g_s = g_{smin} + \frac{1.6 A_{LAI}}{c_a (1 - \lambda_c)}$
LPJ-wsl	Priestley–Taylor	0.5° × 0.5°	Monthly	$g_s = g_{smin} + \frac{1.6 A_{LAI}}{c_a (1 - \lambda_c)}$
LPX-Bern	Modified equation of Monteith (1995)	1° × 1°	Monthly	$g_s = g_{smin} + \frac{1.6 A_{LAI}}{c_a (1 - \lambda_c)}$
O-CN	Modified Penman–Monteith	1° × 1°	Monthly	$g_s = g_{smin} + \frac{1.6 A_{LAI}}{c_a (1 - \lambda_c)}$
ORCHIDEE	Modified Penman–Monteith	0.5° × 0.5°	Monthly	$g_s = g_0 + \frac{A + R_d}{c_a - c_p} f_{vpd}$
ORCHIDEE-MICT	Modified Penman–Monteith	0.5° × 0.5°	Monthly	$g_s = g_0 + \frac{A + R_d}{c_a - c_p} f_{vpd}$
VISIT	Penman–Monteith	0.5° × 0.5°	Monthly	$g_s = g_0 + \frac{g_1 f_{sc} A}{c_a - c_p} \left(1 + \frac{VPD}{VPD_0}\right)^{-1}$
				Io (2010)

Notes:  $A$ : net assimilation rate;  $A_{net}$ : total daytime net photosynthesis;  $A_{unpot}$ : unstressed net assimilation rate;  $b$ : soil moisture factor;  $b_s$ : stomatal conductance intercept;  $c_a$ : atmospheric CO<sub>2</sub> concentration;  $c_c$ : critical CO<sub>2</sub> concentration;  $c_l$ : internal leaf concentration of CO<sub>2</sub>;  $c_{i,pot}$ : internal leaf concentration of CO<sub>2</sub> for unstressed conditions;  $c_s$ : leaf surface CO<sub>2</sub> concentration;  $c_p$ : CO<sub>2</sub> compensation point;  $c_s$ : vapor pressure at leaf surface;  $c_i$ : saturation vapor pressure inside the leaf;  $E_s$ : soil evaporation;  $E_c$ : canopy evapotranspiration;  $E_{dry}$ : dry canopy evapotranspiration;  $E_w$ : wet canopy evapotranspiration;  $E_i$ : canopy transpiration;  $P_{canopy}$ : potential canopy evaporation;  $q_a$ : specific air humidity;  $q_c$ : critical humidity deficit;  $q_s$ : specific humidity of saturated air;  $r_a$ : aerodynamic resistance;  $r_s$ : stomatal resistance;  $R_n$ : net radiation;  $R_d$ : day respiration;  $RH$ : relative humidity;  $T_s$ : actual surface temperature;  $VPD$ : vapor pressure deficit;  $VPD_0$ : the sensitivity of stomatal conductance to VPD;  $W$ : top soil moisture;  $W_{canopy}$ : canopy water;  $W_{sat}$ : soil porosity;  $\alpha$ : Priestley–Taylor coefficient;  $\alpha_{m1}$ : empirical parameter;  $\beta$ : a constant accounting for the times when vegetation is wet;  $\beta_i$ : soil water availability factor between 0 and 1;  $\beta_w$ : a scaling factor to account for water stress;  $\beta_s$ : moisture availability function;  $\rho$ : air density;  $\gamma$ : psychrometric constant;  $\lambda_v$ : latent heat of vaporization;  $\lambda_c$ : ratio of intercellular to ambient partial pressure of CO<sub>2</sub>;  $r_{corb}$ : correction factor of temperature and air pressure on conductance;  $\Gamma^*$ : CO<sub>2</sub> compensation point when leaf day respiration is zero;  $\theta_i$ : parameter of moisture concentration in the top soil layer;  $\theta_c$ : parameter of moisture concentration in the spatially varying critical soil moisture;  $\Delta$ : slope of the vapor pressure curve.

Imaging Spectroradiometer (MODIS) and the PML-CSIRO (Penman–Monteith–Leuning), and two machine-learning datasets, which are based on random forest (RF) and model tree ensemble (MTE), are used in our study. Both machine-learning and physically based remote sensing datasets (six datasets in total) were considered as benchmark products. The ensemble mean of the benchmark products was calculated as the mean value of all machine-learning and physically based satellite estimates as we treated each benchmark dataset equally.

Three of the four remote-sensing-based physical models quantify ET using PM approaches. P-LSH adopts a modified PM approach coupled with biome-specific canopy conductance determined from NDVI (Zhang et al., 2010). The modified P-LSH model used in this study also accounts for the influences of atmospheric CO<sub>2</sub> concentrations and wind speed on canopy stomatal conductance and aerodynamic conductance (Zhang et al., 2015). The MODIS ET model is based on the algorithm proposed by Cleugh et al. (2007). Mu et al. (2007) improved the model performance through the inclusion of environmental stress into the canopy conductance calculation and by explicitly accounting for soil evaporation by combining the complementary relationship hypothesis with the PM equation. The MODIS ET product (MOD16A3) used in this study was further improved by considering nighttime ET, simplifying the vegetation cover fraction calculation, adding the soil heat flux item, dividing the saturated wet and moist soil, separating the dry and wet canopy, as well as modifying algorithms of aerodynamic resistance, stomatal conductance and boundary layer resistance (Mu et al., 2011). PML-CSIRO adopts the Penman–Monteith–Leuning algorithm, which calculates surface conductance and canopy conductance using a biophysical model instead of classic empirical models. The maximum stomatal conductance is estimated using the trial-and-error method (Y. Zhang et al., 2016). Furthermore, for each grid covered by natural vegetation, the PML-CSIRO model constrains ET at the annual scale using the Budyko hydrometeorological model proposed by Fu (1981). The GLEAM ET calculation is based on the PT equation, which requires fewer model inputs than the PM equation, and the majority of these inputs can be directly achieved from satellite observations. Its rationale is to make the most of information about evaporation contained in the satellite-based environmental and climatic observations (Martens et al., 2017; Miralles et al., 2011). Key variables including air temperature, land surface temperature, precipitation, soil moisture, vegetation optical depth and snow water equivalent are satellite-observed values. Moreover, the extensive usage of microwave remote sensing products in GLEAM ensures the accurate estimation of ET under diverse weather conditions. Here, we use GLEAM v3.2, which has overall better quality than previous versions (Martens et al., 2017).

The first used machine-learning model, MTE, is based on the Tree Induction Algorithm (TRIAL) and Evolving Trees with Random Growth (ERROR) algorithm (Jung et al.,

2009). The TRIAL grows model trees from the root node and splits at each node with the criterion of minimizing the sum of squared errors of multiple regressions in both subdomains. ERROR is used to select the model trees that are independent of one another and have the best performance based on the Schwarz criterion. The canopy fraction of absorbed photosynthetic active radiation ( $f_{\text{APAR}}$ ), temperature, precipitation, relative humidity, sunshine hours and potential radiation are used as explanatory variables to train MTE (Jung et al., 2011). The second machine-learning model is the random forest (RF) algorithm, whose rationale is generating a set of independent regression trees by randomly selecting training samples automatically (Breiman, 2001). Each regression tree is constructed using samples selected by a bootstrap sampling method. After fixing the individual tree in entity, the final result is determined by simple averaging. One merit of the RF algorithm is its capability to handle complicated non-linear problems and high dimensional data (Xu et al., 2018). For the RF product used in this study, multiple explanatory variables including the enhanced vegetation index,  $f_{\text{APAR}}$ , the leaf area index, daytime and nighttime land surface temperature, incoming radiation, top-of-atmosphere potential radiation, the index of water availability and relative humidity were used to train regression trees (Bodesheim et al., 2018).

The 14 LSM-derived ET products were from the Trends and Drivers of the Regional Scale Sources and Sinks of Carbon Dioxide (TRENDY) project (including CABLE, CLASS-CTEM, CLM45, DLEM, ISAM, JSBACH, JULES, LPJ-GUESS, LPJ-wsl, LPX-Bern, O-CN, ORCHIDEE, ORCHIDEE-MICT and VISIT). Daily gridded meteorological reanalyses from the CRU-NCEPv8 dataset (temperature, precipitation, long- and shortwave incoming radiation, wind speed, humidity and air pressure) were used to drive the LSMs. The TRENDY simulations were performed in year 2017 and contributed to the global carbon budget reported in Le Quéré et al. (2018). We used the results from the S3 simulation from TRENDYv6 (with changing CO<sub>2</sub>, climate and land use) over the period from 1982 to 2011, which is a time period consistent with other products derived from remote-sensing-based physical models and machine-learning algorithms.

### 2.3 Description of other datasets

To quantify the contributions of vegetation greening to terrestrial ET variations, we used the LAI from the TRENDYv6 S3 simulation. We also used the newest version of the Global Inventory Modeling and Mapping Studies LAI data (GIMMS LAI3gV1) as the satellite-derived LAI. GIMMS LAI3gV1 was generated from AVHRR GIMMS NDVI3g using a model derived from an artificial neural network (ANN; Zhu et al., 2013). It covers the period from 1982 to 2016 with bimonthly frequency and has a 1/12° spatial resolution. To achieve a uniform resolution, all data were resampled to 1/2° using the nearest neighbor method. Follow-

ing N. Pan et al. (2018), grids with an annual mean NDVI of less than 0.1 were assumed to be non-vegetated regions and were, therefore, masked out. NDVI data are from the GIMMS NDVI3gV1 dataset. Temperature, precipitation and radiation are from the CRU-NCEPv8 dataset.

## 2.4 Statistical analysis

The significance of ET trends was analyzed using the Mann-Kendall (MK) test (Kendall, 1955; Mann, 1945). It is a rank-based nonparametric method that has been widely applied for detecting trends in hydro-climatic time series (Sayemuzaman and Jha, 2014; Yue et al., 2002). The Theil–Sen estimator was applied to estimate the magnitude of the slope. The advantage of this method over an ordinary least squares estimator is that it limits the influence of the outliers on the slope (Sen, 1968).

Terrestrial ET interannual variability (IAV) is mainly controlled by variations in temperature, precipitation and short-wave solar radiation (Zeng et al., 2018b; Zhang et al., 2015). In this study, we performed partial correlation analyses between ET and these three climatic variables at an annual scale for each grid cell to explore climatic controls on ET IAV. Variability caused by climatic variables was assessed through the square of partial correlation coefficients between ET and temperature, precipitation and radiation. We chose partial correlation analysis because it can quantify the linkage between ET and a single environmental driving factor while controlling the effects of the other remaining environmental factors. Partial correlation analysis is a widely applied statistical tool to isolate the relationship between two variables from the confounding effects of many correlated variables (Anav et al., 2015; Jung et al., 2017; Peng et al., 2013). All variables were first detrended in the statistical correlation analysis, as we focus on the interannual relationship. The study period is from 1982 to 2011 for all models except MODIS and random forest, which have a temporal coverage that is limited to 2001–2011 because of data availability.

To quantify the contribution of vegetation greening to terrestrial ET, we separated the trend in terrestrial ET into four components induced by climatic variables and vegetation dynamics by establishing a multiple linear regression model between global ET and temperature, precipitation, shortwave radiation and LAI (Eqs. 3, 4):

$$\delta(\text{ET}) = \frac{\partial(\text{ET})}{\partial(\text{LAI})}\delta(\text{LAI}) + \frac{\partial(\text{ET})}{\partial T}\delta(T) + \frac{\partial(\text{ET})}{\partial(P)}\delta(P) + \frac{\partial(\text{ET})}{\partial R}\delta(R) + \varepsilon \quad (3)$$

$$\delta(\text{ET}) = \gamma_{\text{ET}}^{\text{LAI}}\delta\text{LAI} + \gamma_{\text{ET}}^T\delta T + \gamma_{\text{ET}}^P\delta P + \gamma_{\text{ET}}^R\delta R + \varepsilon, \quad (4)$$

where  $\gamma_{\text{ET}}^{\text{LAI}}$ ,  $\gamma_{\text{ET}}^T$ ,  $\gamma_{\text{ET}}^P$  and  $\gamma_{\text{ET}}^R$  are the sensitivities of ET to the leaf area index (LAI), air temperature ( $T$ ), precipitation ( $P$ ) and radiation ( $R$ ), respectively.  $\varepsilon$  is the residual, representing the impacts of other factors.

After calculating  $\gamma_{\text{ET}}^{\text{LAI}}$ ,  $\gamma_{\text{ET}}^T$ ,  $\gamma_{\text{ET}}^P$  and  $\gamma_{\text{ET}}^R$ , the contribution of the trend in factor  $i$  ( $\text{Trend}(i)$ ) to the trend in ET ( $\text{Trend}(\text{ET})$ ) can be quantified as follows:

$$\text{Contri}(i) = (\gamma_{\text{ET}}^i \times \text{Trend}(i)) / \text{Trend}(\text{ET}). \quad (5)$$

In performing multiple linear regression, we used the GIMMS LAI for both remote-sensing-based physical models and machine-learning methods as well as the individual TRENDYv6 LAI for each TRENDY model. The gridded data of temperature, precipitation and radiation are from CRU-NCEPv8.

## 3 Results

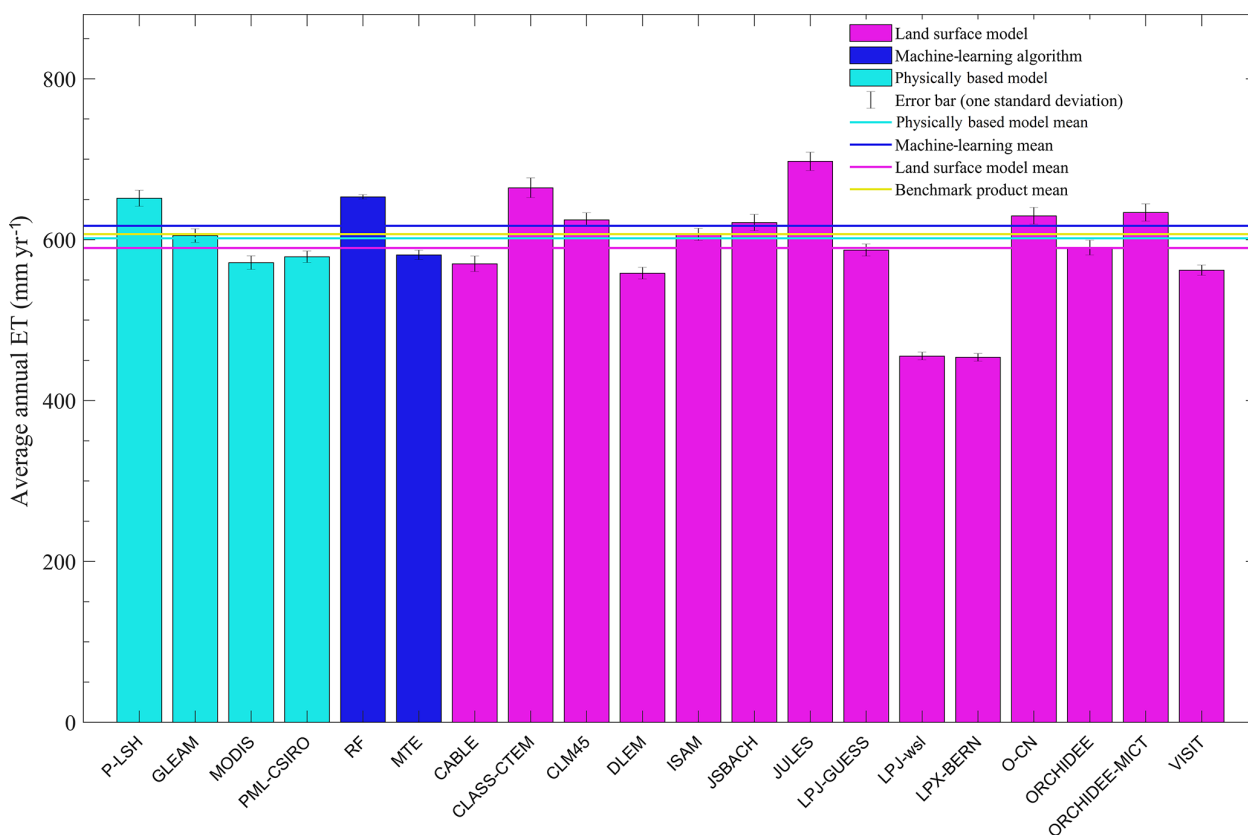
### 3.1 The ET magnitude estimated by multiple models

The multiyear ensemble mean of annual global terrestrial ET from 2001 to 2011 derived by machine-learning methods, remote-sensing-based physical models and the TRENDY models agreed well, with values ranging from 589.6 to 617.1 mm yr<sup>−1</sup>. However, substantial differences existed among individual models (Fig. 1). LPJ-wsl (455.3 mm yr<sup>−1</sup>) and LPX-Bern (453.7 mm yr<sup>−1</sup>) estimated significantly lower ET than other models, even in comparison with most previous studies focusing on earlier periods (Table S1 in the Supplement). On the contrary, JULES gave the largest ET estimate (697.3 mm yr<sup>−1</sup>, which is equal to  $7.57 \times 10^4 \text{ km}^3 \text{ yr}^{-1}$ ) among all models, and showed an obvious increase in ET compared with its estimation from 1950 to 2000 ( $6.5 \times 10^4 \text{ km}^3 \text{ yr}^{-1}$ , Table S1).

### 3.2 Spatial patterns of global terrestrial ET

As shown in Fig. 2, the spatial patterns of the multiyear average annual ET of different categories were similar. ET was highest in the tropics and low in the northern high latitudes and arid regions such as Australia, central Asia, the western US and the Sahel. Compared with remote-sensing-based physical models and LSMs, machine-learning methods obtained a smaller spatial gradient. In general, latitudinal profiles of ET estimated using different approaches were also consistent (Fig. 3). However, machine-learning methods gave higher ET estimate at high latitudes and lower ET in the tropics compared with other approaches. In the tropics, LSMs have significantly larger uncertainties than benchmark products, and the standard deviation of LSMs is about 2 times higher than that of benchmark products (Fig. 3). At other latitudes, LSMs and benchmark ET products generally have comparable uncertainties. The largest difference in ET in the different categories was found in the Amazon Basin (Fig. 2). In most regions of the Amazon Basin, the mean ET of remote sensing physical models is more than 200 mm yr<sup>−1</sup> higher than the mean ET of LSMs and machine-learning methods. For individual ET estimates, the largest uncertainty





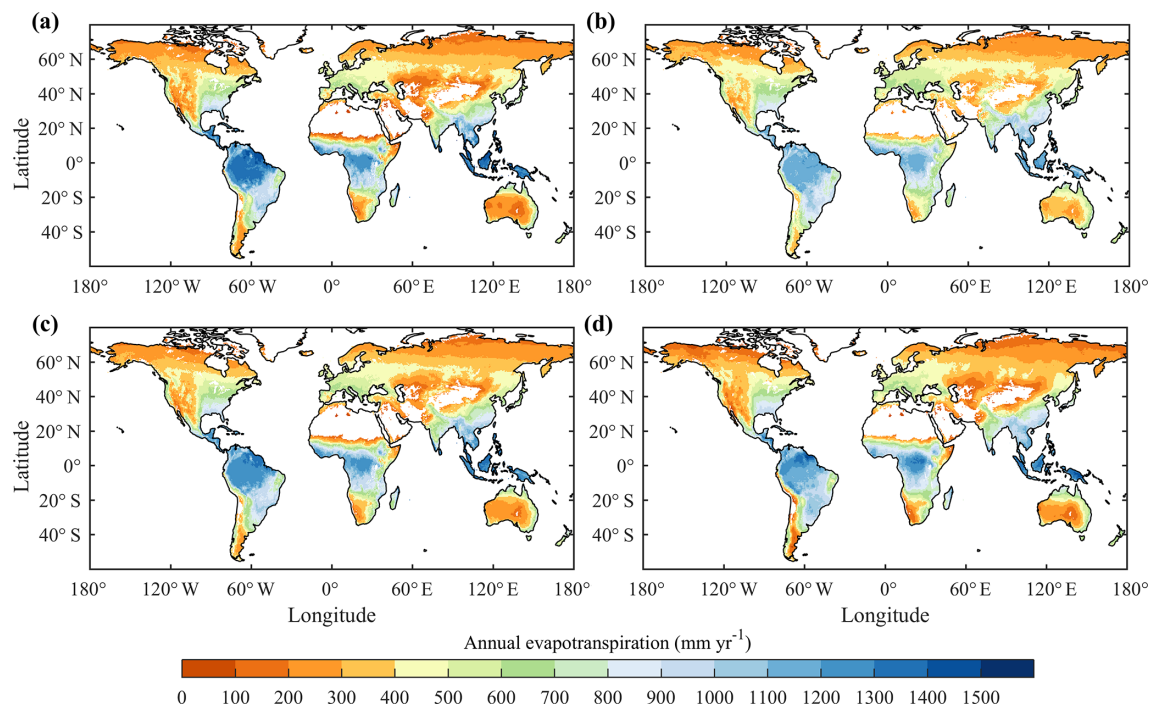
**Figure 1.** Average annual global terrestrial ET estimated by each model during the period from 2001 to 2011. Error bars represent the standard deviation of each model. The four lines indicate the mean value of each category.

was also found in the Amazon Basin. MODIS, VISIT and CLASS-CTEM estimated that the annual ET was higher than 1300 mm in the majority of Amazon, whereas JSBACH and LPJ-wsl estimated an ET value lower than 800 mm yr<sup>-1</sup> (Fig. S1). As is shown in Fig. S2, the differences in ET estimates among TRENDY models were larger than those among benchmark estimates for tropical and humid regions. The uncertainty of ET estimates from LSMs is particularly large in the Amazon Basin, where the standard deviation of LSM estimates is more than 2 times larger than that of benchmark estimates. It is noteworthy that, in arid and semiarid regions such as western Australia, central Asia, northern China and the western US, the difference in ET estimates among LSMs is significantly smaller than those among remote sensing models and machine-learning algorithms.

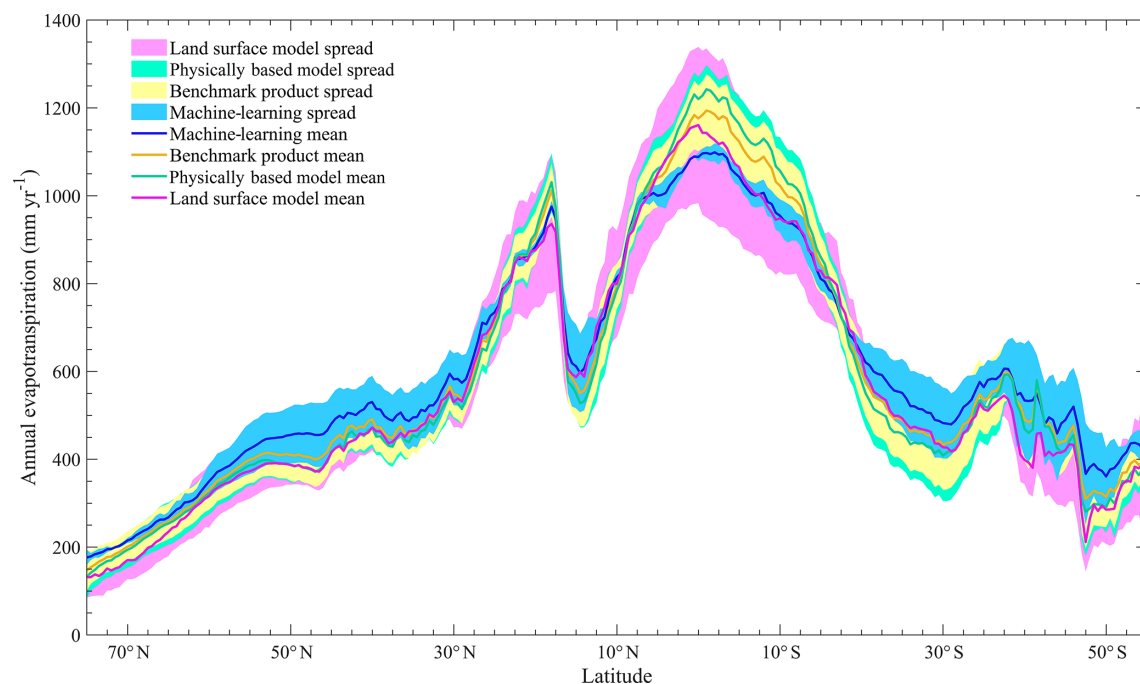
### 3.3 Interannual variations in global terrestrial ET

The ensemble mean interannual variability (IAV) of remote sensing ET estimates and LSM ET estimates showed similar spatial patterns (Fig. 4). Both remote sensing physical models and LSMs presented low IAV in ET in the northern high latitudes but high IAV in ET in the southwestern US, India, sub-Saharan Africa, the Amazon and Australia. In

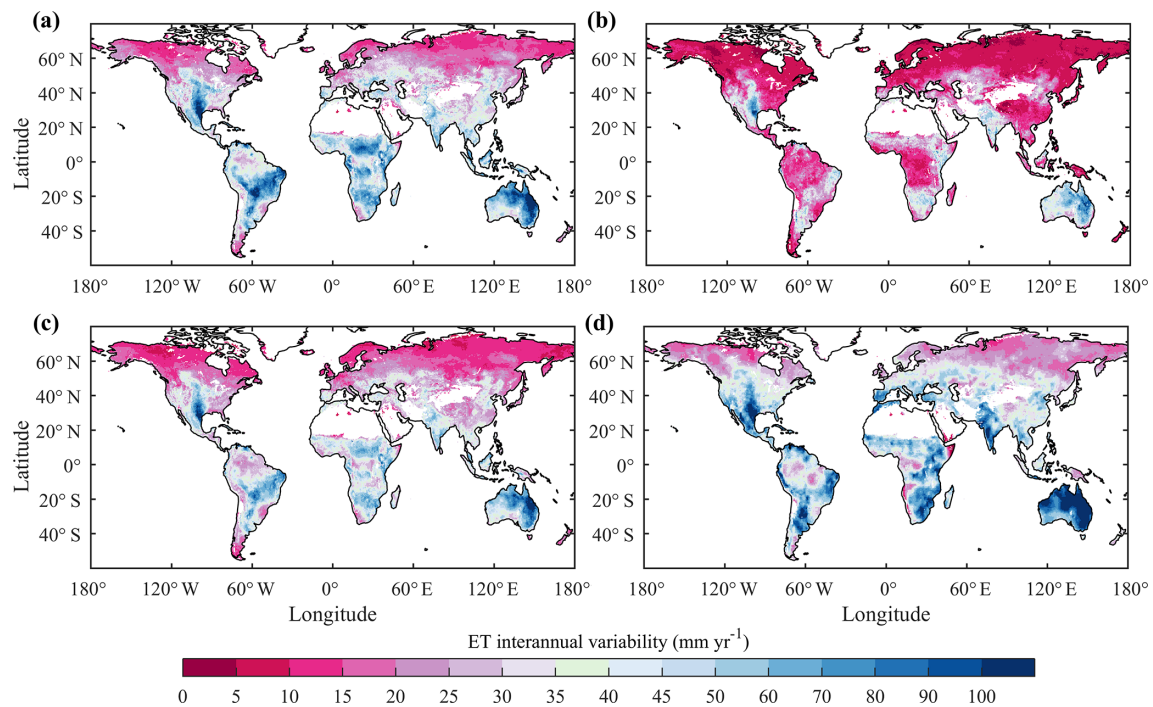
contrast, the IAV of machine-learning-based ET was much weaker. In most regions, the IAV of machine-learning ET is lower than 40 % of the IAV of remote sensing physical ET and LSM ET, and this phenomenon is especially pronounced in tropical regions. Further investigation into the spatial patterns of ET IAV for individual models showed that the two machine-learning methods performed equally with respect to estimating spatial patterns of ET IAV (Fig. S4). In contrast, differences in ET IAV among remote sensing physical estimates and LSM estimates were much larger. LSMs showed the largest differences in the IAV of ET in tropical regions. For example, CABLE and JULES obtained an ET IAV of less than 15 mm yr<sup>-1</sup> in most regions of the Amazon Basin, whereas LPJ-GUESS predicted an ET IAV of more than 60 mm yr<sup>-1</sup>. Figure 5 shows that remote sensing physical ET and LSM ET had comparable IAV north of 20° S, but the IAV of the machine-learning-based ET was much lower in this region. In the region south of 20° S, TRENDY ET showed the largest IAV, followed by those of remote sensing physical ET and machine-learning estimates. The three approaches agreed on that ET IAV in the Southern Hemisphere was generally larger than that in the Northern Hemisphere.



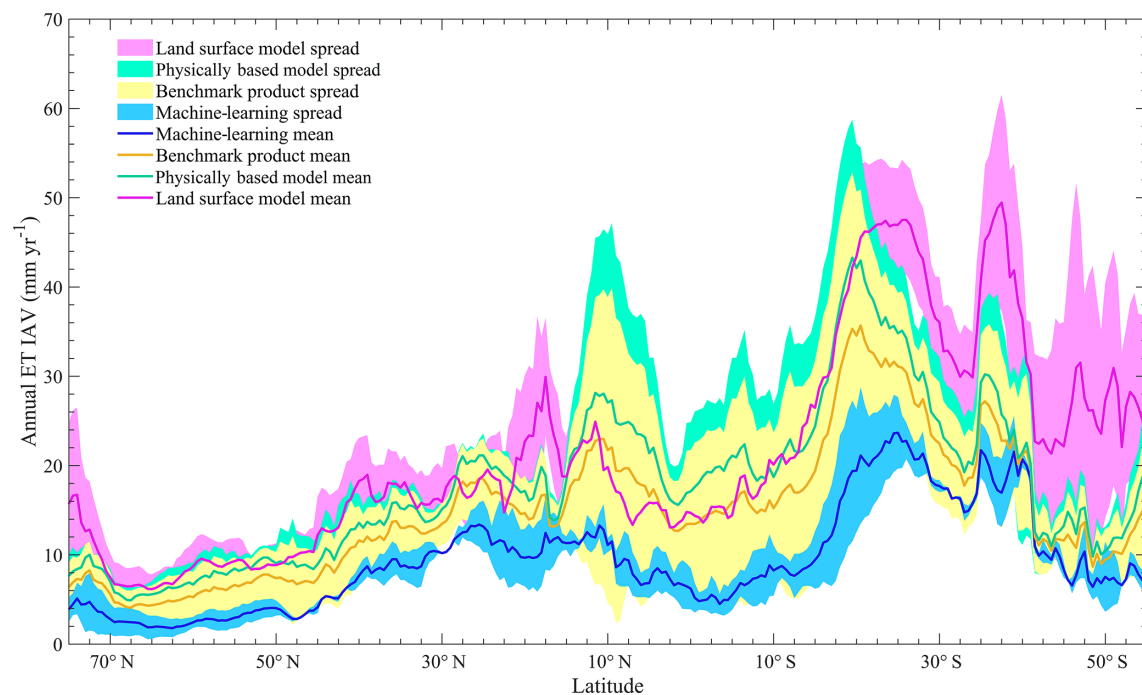
**Figure 2.** Spatial distributions of the mean annual ET derived from (a) remote-sensing-based physical models, (b) machine-learning algorithms, (c) benchmark datasets and (d) the TRENDY LSMs' ensemble mean, respectively.



**Figure 3.** Latitudinal profiles of the mean annual ET for different categories of models. Each line represents the mean value of the corresponding category, and the shading represents the interval of 1 standard deviation.



**Figure 4.** Spatial distributions of the interannual variability in ET derived from (a) remote-sensing-based physical models, (b) machine-learning algorithms, (c) benchmark datasets and (d) the TRENDY LSMs' ensemble mean, respectively. The study period used in this study for the interannual variability analysis is from 1982 to 2011.



**Figure 5.** Latitudinal profiles of ET IAV for different categories of models. Each line represents the mean value of the corresponding category, and the shading represents the interval of 1 standard deviation.

### 3.4 Climatic controls on ET

According to the ensemble remote sensing models, temperature and radiation dominated ET IAV in northern Eurasia, northern and eastern North America, southern China, the Congo River basin, and the southern Amazon River basin, while precipitation dominated ET IAV in arid regions and semiarid regions (Fig. 6a). The ensemble machine-learning algorithms had a similar pattern, but they suggested a stronger control of radiation in the Amazon Basin and a weaker control of precipitation in several arid regions such as central Asia and northern Australia (Fig. 6b). In comparison, the ensemble LSMs suggested the strongest control of precipitation on ET IAV (Fig. 6). According to the ensemble LSMs, ET IAV was dominated by precipitation IAV in most regions of the Southern Hemisphere and at northern low latitudes. Temperature and radiation only controlled northern Eurasia, eastern Canada and part of the Amazon Basin (Fig. 6d). As is shown in Fig. S6, the majority of LSMs agreed on the dominant role of precipitation in controlling ET in regions south of 40° N. However, the pattern of climatic controls in the ORCHIDEE-MICT model is quite unique and different from all of the other LSMs. According to the ORCHIDEE-MICT model, radiation and temperature dominate ET IAV in more regions, and precipitation only controls ET IAV in eastern Brazil, northern Russia, central Europe and a part of tropical Africa. As ORCHIDEE-MICT was developed from ORCHIDEE, the dynamic root parameterization in ORCHIDEE-MICT may explain why ET is less driven by precipitation compared with ORCHIDEE (Haverd et al., 2018). It is noted that two machine-learning algorithms, MTE and RF, showed significant discrepancies in the spatial pattern of dominant climatic factors. According to the result from MTE, temperature controlled ET IAV in regions north of 45° N, the eastern US, southern China and the Amazon Basin (Fig. S6e). By contrast, RF suggested that precipitation and radiation dominated ET IAV in these regions (Fig. S6f).

### 3.5 Long-term trends in global terrestrial ET

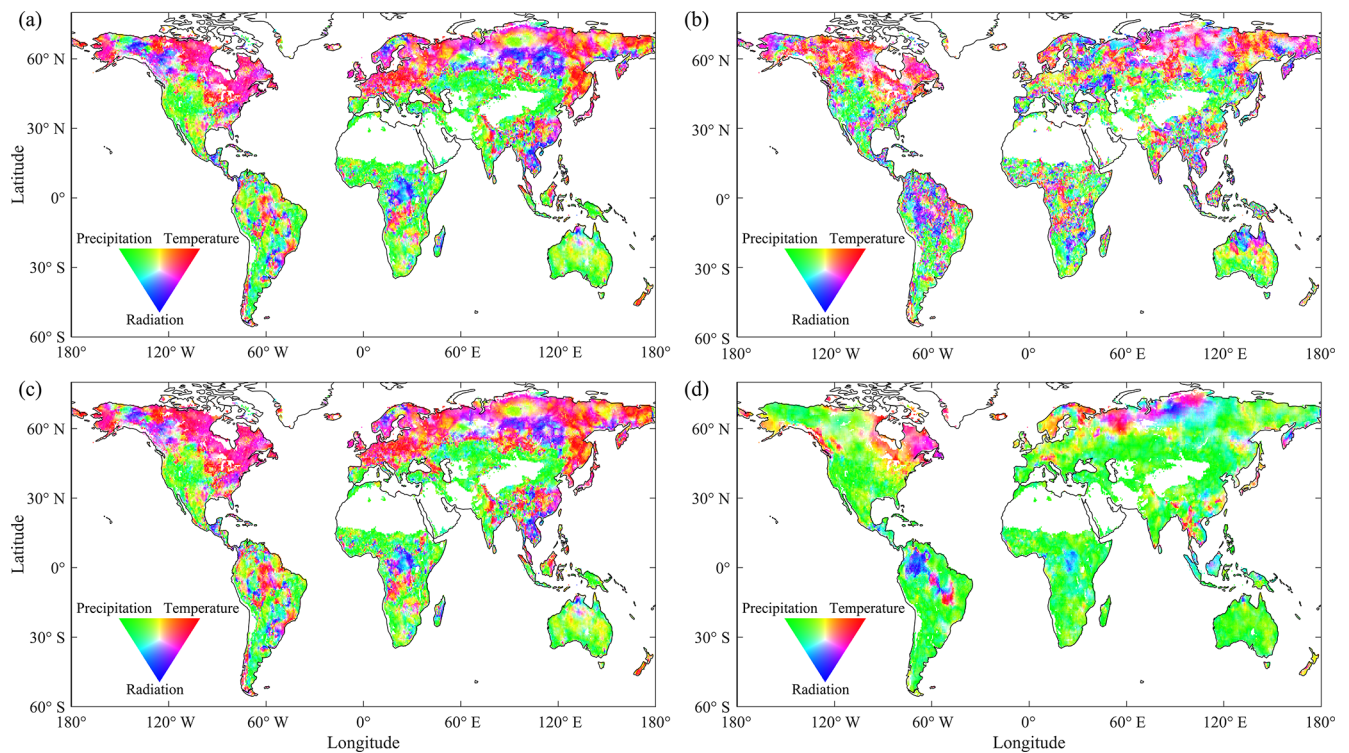
All approaches suggested an overall increasing trend in global ET during the period from 1982 to 2011 (Fig. 7), although ET decreased from 1998 to 2009. This result is consistent with previous studies (Jung et al., 2010; Lian et al., 2018; Zhang et al., 2015). Remote sensing physical models indicated the largest increase in ET ( $0.62 \text{ mm yr}^{-2}$ ), followed by the machine-learning method ( $0.38 \text{ mm yr}^{-2}$ ) and land surface models ( $0.23 \text{ mm yr}^{-2}$ ). The mean ET of all categories except LSMs significantly increased during the study period ( $p < 0.05$ ). It is noted that the ensemble mean ET values of different categories are statistically correlated with each other ( $p < 0.001$ ), even if the driving forces of different ET approaches are different.

All remote sensing and machine-learning estimates indicate a significant increasing trend in ET during the study period ( $p < 0.05$ ), although the increase rate of P-LSH ( $1.07 \text{ mm yr}^{-2}$ ) is more than 3 times as large as that of GLEAM ( $0.33 \text{ mm yr}^{-2}$ ). Nevertheless, there is a larger discrepancy among LSMs in terms of the ET trend. The majority of LSMs (10 of 14) suggest an increasing trend with an average trend of  $0.34 \text{ mm yr}^{-2}$  ( $p < 0.05$ ), and eight of them are statistically significant (see Table 2). However, four LSMs (JSBACH, JULES, ORCHIDEE and ORCHIDEE-MICT) suggest a decreasing trend with an average trend of  $-0.12 \text{ mm yr}^{-2}$  ( $p > 0.05$ ). Among the four decreasing trends, only the trend of ORCHIDEE-MICT ( $-0.34 \text{ mm yr}^{-2}$ ) is statistically significant ( $p < 0.05$ ).

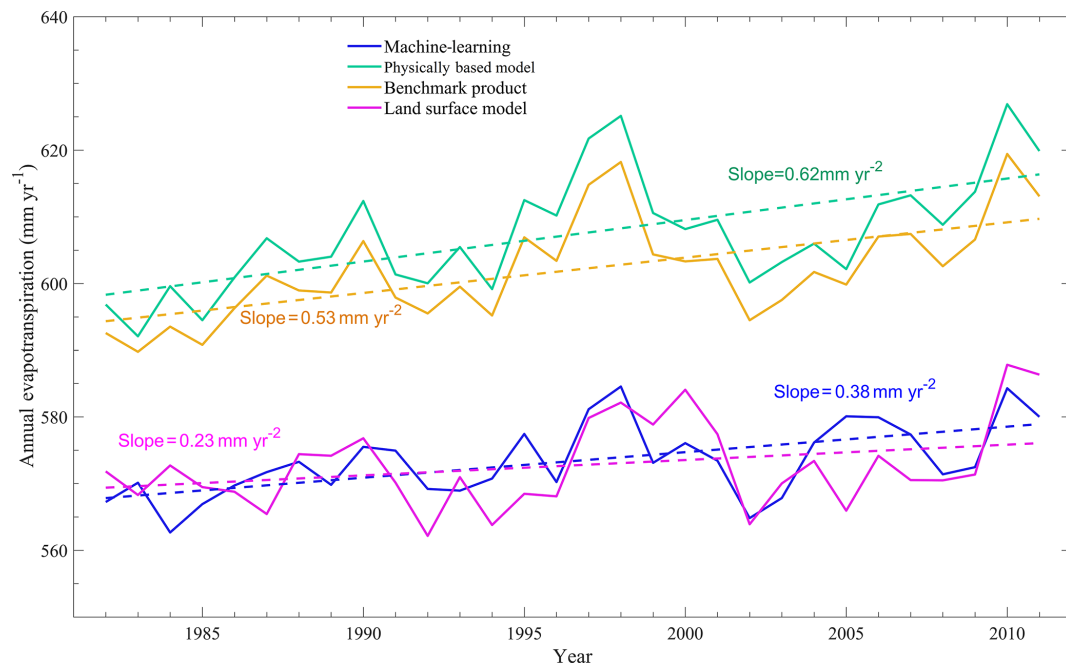
According to Fig. 8, the ensemble means of all the three approaches showed increasing trends in ET over western and southern Africa, western India and northern Australia, and decreasing ET over the western US, southern South America and Mongolia. Discrepancies in ET trends mainly appeared in eastern Europe, eastern India and central China. LSMs also suggested a larger area of decreasing ET in both North America and South America. Although the differences in ET trends among individual models were larger than those among the ensemble means of different approaches, the majority of models agreed that ET increased in western and southern Africa, and decreased in the western US and southern South America (Fig. S2). For both remote sensing estimates and LSM estimates, ET trends in the Amazon Basin had large uncertainty: P-LSH, CLM45 and VISIT suggested a large area of increasing ET, whereas GLEAM, JSBACH and ORCHIDEE suggested a large area of decreasing ET.

### 3.6 Impacts of vegetation changes on ET variations

During the period from 1982 to 2011, global LAI trends estimated from remote sensing data and from the ensemble LSMs are  $2.51 \times 10^{-3} \text{ m}^2 \text{ m}^{-2} \text{ yr}^{-1}$  ( $p < 0.01$ ) and  $4.63 \times 10^{-3} \text{ m}^2 \text{ m}^{-2} \text{ yr}^{-1}$  ( $p < 0.01$ ), respectively (Table 2). All LSMs suggested a significant increasing trend in global LAI (greening). For both benchmark estimates and LSM estimates, it was found that the spatial pattern of trends in ET matched well with that of trends in the LAI (Figs. 8c, d, S5a, b), indicating significant effects of vegetation dynamics on ET variations. According to the results of the multiple linear regression, all models agreed that greening of the Earth since the early 1980s intensified terrestrial ET (Table 2), although there was a significant discrepancy in the magnitude of ET intensification which varied from  $0.04$  to  $0.70 \text{ mm yr}^{-2}$ . The ensemble LSMs suggested a smaller ET increase ( $0.23 \text{ mm yr}^{-2}$ ) than the ensemble remote sensing physical models ( $0.62 \text{ mm yr}^{-2}$ ) and machine-learning algorithm ( $0.38 \text{ mm yr}^{-2}$ ). Nevertheless, the greening-induced ET intensification estimated by LSMs ( $0.37 \text{ mm yr}^{-2}$ ) is larger than that estimated by remote sensing models ( $0.28 \text{ mm yr}^{-2}$ ) and machine-learning algo-



**Figure 6.** Spatial distributions of climatic controls on the interannual variation of ET derived from the ensemble means of remote-sensing-based physical models (a), machine-learning algorithms (b), benchmark data (c) and the TRENDY LSMs (d). (Red denotes temperature, green denotes precipitation and blue denotes radiation.)



**Figure 7.** Interannual variations in global terrestrial ET estimated using different categories of approaches.

**Table 2.** Interannual variability (IAV – denoted as standard deviation) and the trend of global terrestrial ET from 1982 to 2011 and the contribution of vegetation greening to the ET trend. (RS refers to remote sensing.)

	Model	ET IAV (mm yr <sup>-1</sup> )	ET trend (mm yr <sup>-2</sup> )	Greening-induced ET change (mm yr <sup>-2</sup> )	Sensitivity of ET to LAI (mm yr <sup>-2</sup> m <sup>-2</sup> m <sup>-2</sup> )	LAI trend (10 <sup>-3</sup> m <sup>2</sup> m <sup>-2</sup> yr <sup>-1</sup> )
Machine learning	MTE	5.93	0.38*	0.09	35.86	2.51*
RS models	P-LSH	9.95	1.07*	0.34	135.46	2.51*
	GLEAM	8.47	0.33*	0.14	55.78	2.51*
	PML-CSIRO	7.18	0.41*	0.36	143.43	2.51*
	RS model mean	7.98	0.62*	0.28	111.55	2.51*
LSMs	CABLE	9.63	0.07	0.35	102.64	3.41*
	CLASS-CTEM	12.22	0.35*	0.53	134.52	3.94*
	CLM45	8.68	0.38*	0.31	67.54	4.59*
	DLEM	7.21	0.26*	0.53	200.76	2.64*
	ISAM	7.50	0.22	0.16	32.26	4.96*
	JSBACH	10.12	−0.05	0.50	217.39	2.30*
	JULES	11.33	−0.02	0.34	85.21	3.99*
	LPJ-GUESS	7.48	0.50*	0.28	160.92	1.74*
	LPJ-wsl	4.77	0.24*	0.19	31.56	6.02*
	LPX-Bern	4.80	0.20*	0.04	4.04	9.90*
	O-CN	10.41	0.32*	0.53	89.23	5.94*
	ORCHIDEE	9.28	−0.17	0.21	96.33	2.18*
	ORCHIDEE-MICT	10.70	−0.34*	0.50	171.23	2.92*
	VISIT	6.31	0.87*	0.70	51.40	13.62*
	LSM mean	7.73	0.23	0.37	79.91	4.63*

\* Suggests significance at the 95 % confidence level ( $p < 0.05$ ).

rhythms (0.09 mm yr<sup>-2</sup>), as LSMs suggested a stronger greening trend than remote sensing models. The contribution of vegetation greening to ET intensification estimated by the ensemble LSMs is larger than 100 %, whereas the contributions estimated by the ensemble remote sensing physical models (0.62 mm yr<sup>-2</sup>) and machine-learning algorithm are smaller than 50 %. Although TRENDY LSMs were driven by the same climate data and remote sensing physical models were driven by varied climate data, TRENDY LSMs still showed a larger discrepancy in terms of the effect of vegetation greening on terrestrial ET than remote sensing physical models due to the significant differences in both LAI trends (1.74–13.63 × 10<sup>-3</sup> m<sup>2</sup> m<sup>-2</sup> yr<sup>-1</sup>) and the sensitivities of ET to the LAI (4.04–217.39 mm yr<sup>-2</sup> m<sup>-2</sup> m<sup>-2</sup>). In comparison, remote sensing physical models had smaller discrepancies in terms of the sensitivity of ET to LAI (55.78–143.43 mm yr<sup>-2</sup> m<sup>-2</sup> m<sup>-2</sup>).

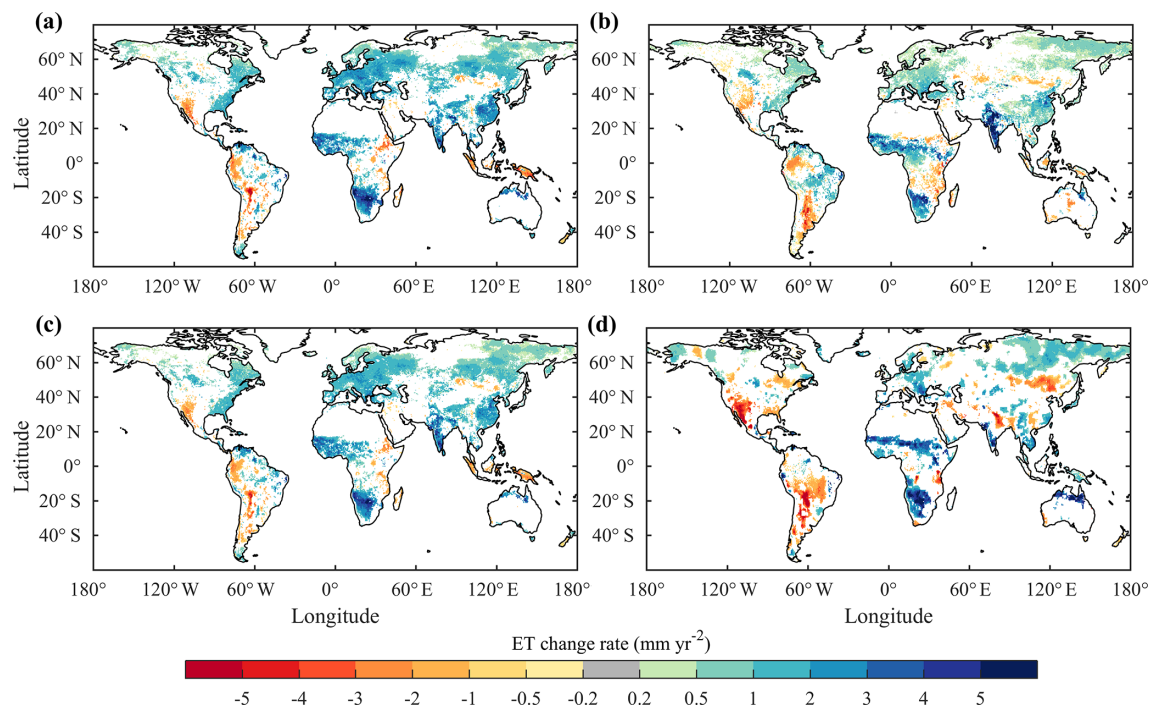
## 4 Discussion and perspectives

### 4.1 Sources of uncertainty

#### 4.1.1 Uncertainty in the ET estimation of the Amazon Basin

LSMs show large discrepancies in the magnitude and trend of ET in the Amazon Basin (Figs. 3, S3). However, it is challenging to identify the uncertainty sources. Given that the TRENDY LSMs used uniform meteorological inputs, the discrepancies in ET estimates among the participating models mainly arise from the differences in the underlying model structures and parameters. One potential source of uncertainty is the parameterization of root water uptake. In the Amazon Basin, a deep root depth has been confirmed using field measurements (Nepstad et al., 2004). However, many LSMs have an unrealistically shallow rooting depth (generally less than 2 m), thereby neglecting the existence and significance of deep roots. The incorrect root distributions enlarge the differences in plant available water and root water uptake, producing large uncertainties in ET. In addition, differences in the parameterization of other key processes pertinent to ET such as LAI dynamics (Fig. S5), canopy conductance variations (Table 1), water movements in the soil





**Figure 8.** Spatial distributions of ET trends from 1982 to 2011 derived from (a) remote-sensing-based physical models, (b) machine-learning algorithms, (c) benchmark datasets and (d) the TRENDY LSMs' ensemble mean, respectively. Regions with nonsignificant trends were excluded.

(Abramopoulos et al., 1988; Clark et al., 2015; Noilhan and Mahfouf, 1996) and soil moisture's control on transpiration (Purdy et al., 2018; Szutu and Papuga, 2019) also increase the uncertainty in ET. The abovementioned processes are not independent of each other but interact in complex ways to produce the end result.

#### 4.1.2 Uncertainty in the ET estimation of arid and semiarid regions

In arid and semiarid regions, benchmark products show much larger differences in the magnitude of ET than LSMs (Fig. S2). One cause of this phenomenon is the difference in meteorological forcing. Remote sensing and machine-learning datasets used different forcing data. For precipitation, RF used the CRUNCEPv6 dataset, MTE used the Global Precipitation Climatology Centre (GPCC) dataset, MODIS used the Global Modeling and Assimilation Office (GMAO) dataset, GLEAM used the Multi-Source Weighted-Ensemble Precipitation (MSWEP) dataset, PML-CSIRO used the Princeton Global Meteorological Forcing (PGF) and the WATCH Forcing Data ERA-Interim (WFDEI) datasets, and P-LSH used data derived from four independent sources. As precipitation is the key climatic factor controlling ET in arid and semiarid regions (Fig. 6), discrepancies between different forcing precipitation (Sun et al., 2018) may be the main source of large uncertainty there. In comparison, the uniform forcing data reduced the inter-model range in ET es-

timates from the TRENDY LSMs. Nevertheless, it is noted that the congruence across LSM ET estimates does not necessarily mean that they are the correct representation of ET. The narrower inter-model range may suggest shared biases. All remote sensing models and machine-learning algorithms except GLEAM do not explicitly take the effects of soil moisture into account (Table S1). Given that soil moisture is pivotal to both canopy conductance and soil evaporation in arid and semiarid regions (A et al., 2019; De Kauwe et al., 2015; Medlyn et al., 2015; Purdy et al., 2018), the lack of soil moisture information also increases the bias in ET estimation. In addition, the accuracy of remote sensing data itself is also an uncertainty source. The retrieval of key land surface variables, such as the leaf area index and surface temperature, is influenced by vegetation architecture, solar zenith angle and satellite observational angle, particularly over heterogeneous surface (Norman and Becker, 1995).

#### 4.1.3 Uncertainty in the ET IAV in the Southern Hemisphere

In regions south of 20° S (including Australia, southern Africa and southern South America), the ET IAV of remote sensing models and machine-learning algorithms is lower than that of LSMs (Figs. 4, 5), although their spatial patterns are similar. In these regions, GLEAM, the only remote sensing model that explicitly considers the effects of soil moisture, has larger ET IAV than other remote sensing models

and has similar ET IAV to LSMs (Fig. S4). This could imply that most of the existing remote sensing models may underestimate ET IAV in the Southern Hemisphere because the effects of soil moisture are not explicitly considered. Machine-learning algorithms show much lower IAV than other models (Figs. 4, S4). The main reason for this is that ET IAV is partly neglected in the training process, as the magnitude of ET IAV is usually smaller than the spatial and seasonal variability (Anav et al., 2015; Jung et al., 2019). Moreover, the IAV of satellite-based key land surface variables such as the LAI,  $f_{\text{APAR}}$  and surface temperature may be not reliable because of the effects of clouds, which, in turn, affects the estimation of the IAV of satellite-based ET. It is noted that LSM ET IAV shows large differences at latitudes south of  $20^\circ\text{S}$  (Fig. 5). This divergence in ET IAV indicates that LSMs require better representation of the ET response to climate in the Southern Hemisphere.

#### 4.1.4 Uncertainty in the global ET trend

All three categories of ET models detected an overall increasing trend in global terrestrial ET since the early 1980s, which is in agreement with previous studies (Mao et al., 2015; Miralles et al., 2014; Zeng et al., 2018a, b, 2014; Zhang et al., 2015; Y. Zhang et al., 2016). Benchmark products generally suggested stronger ET intensification than LSMs. The weaker ET intensification in LSMs may be induced by the response of stomatal conductance to the increasing atmospheric  $\text{CO}_2$  concentration. Increasing  $\text{CO}_2$  affects ET in two ways: on the one hand, increasing  $\text{CO}_2$  can effectively reduce stomatal conductance and, thus, decrease transpiration (Heijmans et al., 2001; Leipprand and Gerten, 2006; Swann et al., 2016); on the other hand, it can increase vegetation productivity and, thus, increase LAI. For benchmarks, the second effect could be captured by remotely sensed LAI, NDVI or  $f_{\text{APAR}}$ , whereas the first effect was neglected by all models except P-LSH (Zhang et al., 2015). In contrast, both effects were modeled in all TRENDY LSMs.

LAI dynamics have significant influences on ET. The increased LAI trend (greening) since the early 1980s has been reported by previous studies (Mao et al., 2016; Zhu et al., 2016) and has also been confirmed by remote sensing data and all of the TRENDY LSMs used in this study (Table 2, Fig. S5). Zhang et al. (2015) found that the increasing trend of global terrestrial ET from 1982 to 2013 was mainly driven by an increase in the LAI and the enhanced atmosphere water demand. Using a land–atmosphere coupled global climate model (GCM), Zeng et al. (2018b) further estimated that global LAI increased by about 8 %, resulting in an increase of  $0.40 \pm 0.08 \text{ mm yr}^{-2}$  in global ET (contributing to  $55 \% \pm 25 \%$  of the ET increase). This number is close to the estimates of ensemble LSMs ( $0.37 \pm 0.18 \text{ mm yr}^{-2}$ ). In comparison, the remote sensing models and machine-learning algorithms used in this study suggested smaller greening-induced ET increases. It is noted that TRENDY LSMs still

showed a larger discrepancy in terms of the effect of vegetation greening on terrestrial ET than remote sensing physical models (Table 2) due to the significant differences in the LAI trend ( $1.74\text{--}13.63 \times 10^{-3} \text{ m}^2 \text{ m}^{-2} \text{ yr}^{-1}$ ) and in the sensitivity of ET to LAI ( $4.04\text{--}217.39 \text{ mm yr}^{-2} \text{ m}^{-2} \text{ m}^{-2}$ ). Uncertainties in the LAI trend may arise from inappropriate carbon allocations and deficiencies in responding to water deficits (Anav et al., 2013; Hu et al., 2018; Murray-Tortarolo et al., 2013; Restrepo-Coupe et al., 2017). Additionally, for machine-learning algorithms, the results from insufficient long-term in situ measurements and sparse observations in tropical, boreal and arid regions imply that there are likely deficiencies in representing the temporal variations.

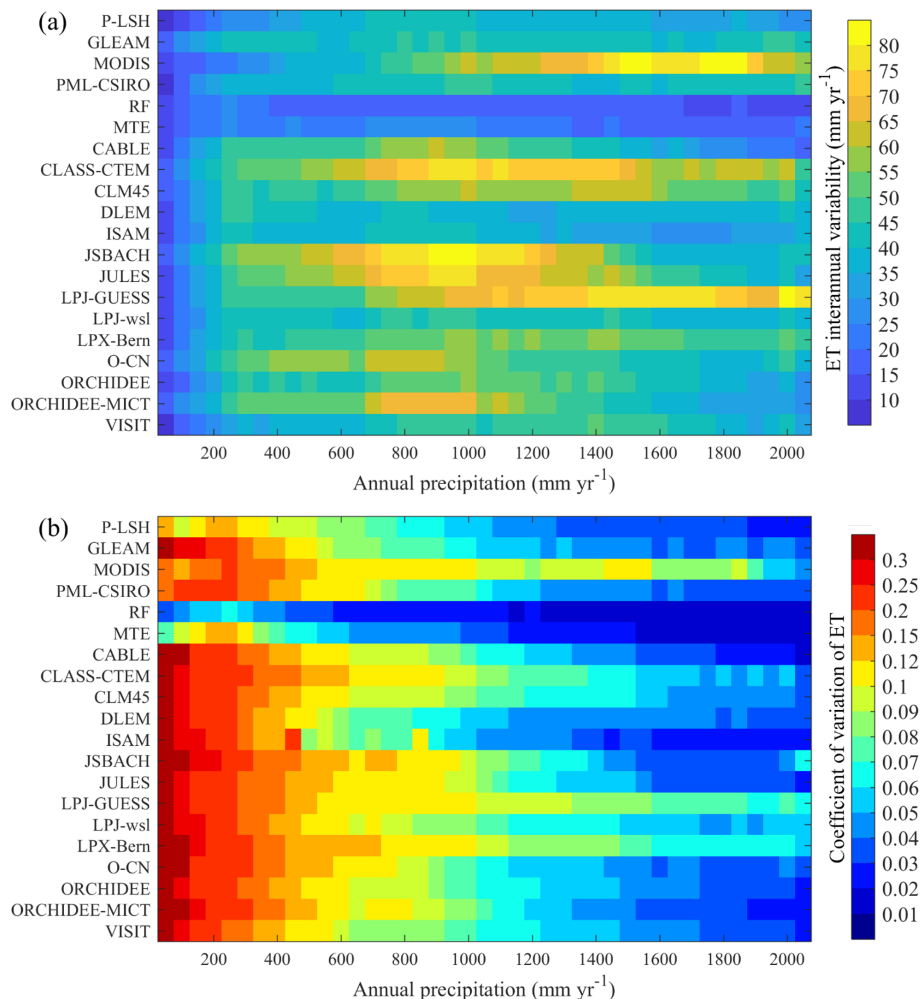
#### 4.1.5 Lack of knowledge of the effects of irrigation

Irrigation accounts for about 90 % of human consumptive water use and largely affects ET in irrigated croplands (Siebert et al., 2010). Global water withdrawals for irrigation were estimated to be within the range of  $1161\text{--}3800 \text{ km}^3 \text{ yr}^{-1}$  around the year 2000, and they largely increased during the period from 2000 to 2014 (Chen et al., 2019). However, none of the remote-sensing-based physical models and machine-learning algorithms explicitly accounted for the effects of irrigation on ET, although these effects could be taken into account to some extent by using observed LAI, NDVI or  $f_{\text{APAR}}$  to drive the models (Zhang et al., 2015). Considering that annual ET may surpass annual precipitation in cropland areas, Y. Zhang et al. (2016) only used the Budyko hydrometeorological model to constrain the PML-CSIRO model in grids covered by non-crop vegetation. However, the process of irrigation affecting evaporation was still not taken into consideration. For TRENDY LSMs, only 2 of 14 models (DLEM and ISAM) included irrigation processes (Le Quéré et al., 2018). Therefore, the effects of irrigation are largely neglected in existing global ET datasets, which reduces the accuracy of local ET estimates in regions with a large proportion of irrigated cropland.

#### 4.1.6 ET variability across the precipitation gradient and its planetary boundary

Precipitation is the source of terrestrial ET. According to Fig. 9a, the vast majority of models agree that ET has the largest IAV in regions with annual precipitation between 700 and 1000 mm, although the magnitude of ET IAV has substantial discrepancies among different models. The low ET IAV in arid and semiarid regions does not mean that ET is stable in these regions. In fact, ET has the largest coefficient of variation (CoV – the ratio of ET standard deviation to ET mean value) in arid regions, and all models show a clear negative trend in the CoV with increasing precipitation (Fig. 9b). This is mainly caused by the large CoV of precipitation in arid regions (Fatichi et al., 2012).





**Figure 9.** Interannual variability (a) and the coefficient of variation (b) of ET for each 50 mm interval of mean annual precipitation.

In comparison, terrestrial ET shows a much lower IAV at the global scale (Table 2), with values ranging from 4.8 to 12.2 mm yr<sup>-1</sup> (1 standard deviation), which only equates to 1.0 %–1.8 % of the global annual mean ET. The model results suggest that global terrestrial ET stabilizes at about  $6.74 \times 10^4 \text{ km}^3 \text{ yr}^{-1}$  (603 mm yr<sup>-1</sup>), which is close to previous estimates (Alton et al., 2009; Mueller et al., 2011; Oki and Kanae, 2006; Zeng et al., 2012). The stability of global terrestrial ET is probably based on partitioning the solar constant and suggests that droughts in one place are balanced by excess rain in other places; thus, it all evens out at a global scale. This implies that ET also has a potential planetary boundary, a suggestion made by Running (2012) on net primary production (NPP) as a planetary boundary. ET integrates four aspects of the current planetary boundaries defined by Steffen et al. (2015): climate change, freshwater use, land-system change and biochemical flows. Given the importance of ET with respect to linking terrestrial water, carbon, nutrient and energy cycles, more studies on the ET planetary boundary are needed in the context of intensifying

global change and increasing anthropogenic perturbations to the Earth system.

In short, the multi-model intercomparison indicates that considerable uncertainty exists in both the temporal and spatial variations in global ET estimates, although a large portion of models adopt similar ET algorithms (Table 1). The major uncertainty source is different for different types of models and regions. The uncertainty is induced by multiple factors, including problems pertinent to parameterization of land processes, lack of in situ measurements, remote sensing acquisition, scaling effects and meteorological forcing. Based on the results of different approaches, we suggest that global terrestrial ET also has a potential planetary boundary, with the value being about  $6.74 \times 10^4 \text{ km}^3 \text{ yr}^{-1}$  (603 mm yr<sup>-1</sup>), which is consistent with previous estimates.

## 4.2 Recommendations for future development

### 4.2.1 Remote-sensing-based physical methods

Over the past few decades, the development of remote sensing technologies has contributed to a boom in various ET estimation methods. However, there is still much room for remote sensing technologies to improve (Fisher et al., 2017). The development of new platforms and sensors that have improved global spatiotemporal coverage and using multiband, multisource remote sensing data are the key points. Planned or newly launched satellites, such as NASA's GRACE Follow-On (GRACE-FO) mission and the ECOSystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) mission, will improve the accuracy of terrestrial ET estimates. ECOSTRESS's thermal infrared (TIR) multispectral scanner is capable of monitoring diurnal temperature patterns at high resolutions, which provides insights into plant response to water stress as well as the means to understand sub-daily ET dynamics (Hulley et al., 2017). GRACE Follow-On observations can be used to constrain subsurface lateral water transfers, which helps to correct soil moisture and subsequently improves the accuracy of ET estimates (Rouholahnejad and Martens, 2018). Moreover, building integrated methods that fuse different ET estimates or the upstream satellite-based biophysical variables from different platforms and the other forcing data will be helpful to improve the accuracy and spatiotemporal coverage of ET (Ke et al., 2016; Ma et al., 2018; Semmens et al., 2016).

The theories and retrieval algorithms of ET and the related key biophysical variables also need to be further improved. For example, the method for calculating canopy conductance may be improved by integrating remote-sensing-based solar-induced chlorophyll fluorescence (SIF) data. SIF data from existing Global Ozone Monitoring Experiment-2 (GOME-2), Orbiting Carbon Observatory-2 (OCO-2) and TROPospheric Monitoring Instrument (TROPOMI) satellite instruments as well as the forthcoming OCO-3 and Geostationary Carbon Cycle Observatory (GeoCarb) satellites provide a good opportunity to diagnose transpiration and to examine ET partitioning at multiple spatiotemporal scales (Pagán et al., 2019; Stoy et al., 2019; Sun et al., 2017). Theoretical advancements in nonequilibrium thermodynamics and maximum entropy production (MEP) could be incorporated into the classical ET theories (Xu et al., 2019; K. Zhang et al., 2016). In addition, quantifying the effects of CO<sub>2</sub> fertilization on stomatal conductance is pivotal for remote sensing models to capture the long-term trend of terrestrial ET.

Most existing remote-sensing-based ET studies have focused on total ET; however, the partitioning of ET between transpiration, soil evaporation, and canopy interception may show significant divergence even though the total ET is accurately estimated (Talsma et al., 2018b). In current remote-sensing-based ET models, soil evaporation, which is sensitive to precipitation events and soil moisture, is the compo-

nent with the largest error (Talsma et al., 2018a). Therefore, incorporating the increasingly accessible satellite-based precipitation, soil moisture observations and soil property data will contribute to the improvement of the soil evaporation estimation. Meanwhile, the consideration of soil evaporation under herbaceous vegetation and canopy will also reduce the errors.

### 4.2.2 Machine-learning methods

It is well known that the capability of machine-learning algorithms to provide accurate ET estimates largely depends on the representativeness of the training datasets with respect to describing ecosystem behaviors (Yao et al., 2017). As a result, machine-learning algorithms may not perform well outside the range of the data used for their training. Unfortunately, long-term field observations in northern temperate regions are still insufficient. This is an important cause of the small spatial gradient and small IAV of machine-learning ET. Given that remote sensing is capable of providing broad coverage of key biophysical variables at reasonable spatial and temporal resolutions, one way to overcome this challenge is to exclusively use remote sensing observations as training data (Jung et al., 2019; Poon and Kinoshita, 2018). Another simple way to make the IAV of machine-learning ET more realistic is to normalize the yearly anomalies when comparing them with ET estimates from LSMs and remote sensing physical models (Jung et al., 2019). New machine-learning techniques, including the extreme learning machine and the adaptive neuro-fuzzy inference system, can be used to improve the accuracy of ET estimation (Gocic et al., 2016; Kişi and Tombul, 2013). Emerging deep-learning methods such as recurrent neural network (RNN) and long short-term memory (LSTM) have the potential to outcompete conventional machine-learning methods in modeling ET time series (Reichstein et al., 2018, 2019). Almost all machine-learning datasets used precipitation rather than soil moisture as the explanatory variable when training. However, soil moisture (rather than precipitation) directly controls ET. As more and more global remote-sensing-based soil moisture datasets become available, using soil moisture products as input is expected to improve the accuracy of ET estimates, especially for regions with sparse vegetation coverage (Xu et al., 2018).

### 4.2.3 Land surface models

In contrast to observation-based methods, LSMs are able to project future changes in ET, and can disentangle the effects of different drivers of ET through factorial analysis. However, results from LSMs are only as good as their parameterizations of complex land surface processes, which are limited by our incomplete understanding of physical and biological processes (Niu et al., 2011). Although TRENDY LSMs are the state-of-the-art process-based global land surfaces models, improvements are still needed because several important

processes are missing or not being appropriately parameterized. Most of the TRENDY LSMs did not simulate processes relevant to human management, including irrigation (Chen et al., 2019) and the application of fertilizers (Mao et al., 2015), or natural disturbances like wildfire (Poon and Kinoshita, 2018). Incorporating these processes into present LSMs is critical, although the introduction of new model parameters also potentially leads to an increase in a model's uncertainty.

In light of the importance of soil water availability in constraining canopy conductance and dynamics, the accurate representation of hydrological processes is a core task for LSMs, particularly in dry regions. Integrating a dynamic root water uptake function and hydraulic redistribution into the LSM can significantly improve its performance with respect to estimating seasonal ET and soil moisture (Li et al., 2012). Moreover, other hydrological processes including groundwater (Decker, 2015), lateral flow (Rouholahnejad and Martens, 2018) and water vapor diffusion at the soil surface (Chang et al., 2018) need to be simulated and correctly represented to reproduce the dynamics of soil water and ET. As the canopy LAI plays an important role in regulating ET, correctly simulating vegetation dynamics is also critical. One way to do this is to correct the initialization, distribution and parameterization of vegetation phenology in LSMs (Murray-Tortarolo et al., 2013; Zhang et al., 2019). An appropriate carbon allocation scheme and parameterization of vegetation's response to water deficits are also important for reproducing vegetation dynamics (Anav et al., 2013).

## 5 Conclusion

In this study, we evaluated 20 global terrestrial ET estimates including 4 from remote-sensing-based physical models, 2 from machine-learning algorithms and 14 from TRENDY LSMs. The ensemble mean values of global terrestrial ET for the three categories agreed well, with values ranging from 589.6 to 617.1 mm yr<sup>-1</sup>. All three categories detected an overall increasing trend in global ET during the period from 1982 to 2011 and suggested a positive effect of vegetation greening on ET intensification. However, the multi-model intercomparison indicates that considerable uncertainties still exist in both temporal and spatial variations in global terrestrial ET estimates. LSMs showed significant differences in the ET magnitude in tropical regions, especially in the Amazon Basin, whereas benchmark ET products showed a larger inter-model range in arid and semiarid regions than LSMs. Trends in ET estimates also showed significant discrepancies among LSMs. These uncertainties are induced by the parameterization of land processes, meteorological forcings, the lack of in situ measurements, remote sensing acquisition and scaling effects. Model developments and observational improvements provide two parallel pathways towards improving the accuracy of global terrestrial ET estimation.

*Code and data availability.* TRENDYv6 data are available from Stephen Sitch (s.a.sitch@exeter.ac.uk) upon reasonable request. MODIS ET data are available from [http://files.ntsg.umd.edu/data/NTSG\\_Products/MOD16/](http://files.ntsg.umd.edu/data/NTSG_Products/MOD16/) (Running, 2020). GLEAM ET are available from <https://www.gleam.eu/> (Miralles, 2020). Both model tree ensemble and random forest ET data are available from <https://www.bgc-jena.mpg.de/geodb/projects/FileDetails.php> (Jung, 2020). P-LSH ET data are available from [http://files.ntsg.umd.edu/data/ET\\_global\\_monthly/Global\\_8kmResolution/](http://files.ntsg.umd.edu/data/ET_global_monthly/Global_8kmResolution/) (K. E. Zhang, 2020). PML-CSIRO ET data are available from <https://data.csiro.au/dap/landingpage?pid=csiro:17375> (Y. Zhang, 2020). CRU-NCEPv8 data are available from Nicolas Viovy upon reasonable request (email address: [viovy@dsm-mail.saclay cea.fr](mailto:viovy@dsm-mail.saclay cea.fr)). GIMMS LAI3gV1 data are available from Ranga B. Myneni upon reasonable request (email address: [rmyneni@bu.edu](mailto:rmyneni@bu.edu)). GIMMS NDVI3gV1 data are available from <https://ecocast.arc.nasa.gov/data/pub/gimms/3g.v1/> (Pinzon and Tucker, 2020).

*Supplement.* The supplement related to this article is available online at: <https://doi.org/10.5194/hess-24-1485-2020-supplement>.

*Author contributions.* SP initiated this research and was responsible for the integrity of the work as a whole. NP carried out the analyses. SP, NP, HT and HS wrote the paper with contributions from all authors. PF, SS, VKA, VH, AKJ, EK, SL, DL, JEMSN, CO, BP, HT and SZ contributed to the TRENDY results. SWR provided the MODIS ET dataset.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* We thank all of the contributors who provided the data used in this study, particularly the TRENDY modeling groups. Additional details on funding support for the participating 14 land surface models is provided by the TRENDY Project.

*Financial support.* This research has been supported by the National Science Foundation (grant nos. 1903722, 1243232 and AGS 12-43071), the US Department of Energy (grant no. DE-SC0016323), the OUC-AU Joint Center program, the Auburn University IGP program, the National Institute of Food and Agriculture/US Department of Agriculture (grant nos. 2015-67003-23489 and 2015-67003-23485), EC H2020 (CCiCC; grant no. 821003), SNSF (grant no. 20020\_172476), and the Earth Systems and Climate Change Hub (funded by the Australian Government's National Environmental Science Program).

*Review statement.* This paper was edited by Pierre Gentile and reviewed by two anonymous referees.

## References

- A, Y., Wang, G., Liu, T., Xue, B., and Kuczera, G.: Spatial variation of correlations between vertical soil water and evapotranspiration and their controlling factors in a semi-arid region, *J. Hydrol.*, 574, 53–63, 2019.
- Abramopoulos, F., Rosenzweig, C., and Choudhury, B.: Improved ground hydrology calculations for global climate models (GCMs): Soil water movement and evapotranspiration, *J. Climate*, 1, 921–941, 1988.
- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration-Guidelines for computing crop water requirements – FAO Irrigation and drainage paper 56, Fao, Rome, Italy, 300, D05109, 1998.
- Alton, P., Fisher, R., Los, S., and Williams, M.: Simulations of global evapotranspiration using semiempirical and mechanistic schemes of plant hydrology, *Global Biogeochem. Cy.*, 23, GB4023, <https://doi.org/10.1029/2009GB003540>, 2009.
- Anav, A., Murray-Tortarolo, G., Friedlingstein, P., Sitch, S., Piao, S., and Zhu, Z.: Evaluation of land surface models in reproducing satellite Derived leaf area index over the high-latitude northern hemisphere. Part II: Earth system models, *Remote Sens.*, 5, 3637–3661, 2013.
- Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., and Peylin, P.: Spatiotemporal patterns of terrestrial gross primary production: A review, *Rev. Geophys.*, 53, 785–818, 2015.
- Antonopoulos, V. Z., Gianniou, S. K., and Antonopoulos, A. V.: Artificial neural networks and empirical equations to estimate daily evaporation: application to lake Vegoritis, Greece, *Hydrolog. Sci. J.*, 61, 2590–2599, 2016.
- Barman, R., Jain, A. K., and Liang, M.: Climate-driven uncertainties in modeling terrestrial energy and water fluxes: a site-level to global-scale analysis, *Glob. Change Biol.*, 20, 1885–1900, 2014.
- Bodesheim, P., Jung, M., Gans, F., Mahecha, M. D., and Reichstein, M.: Upscaled diurnal cycles of land–atmosphere fluxes: a new global half-hourly data product, *Earth Syst. Sci. Data*, 10, 1327–1365, <https://doi.org/10.5194/essd-10-1327-2018>, 2018.
- Breiman, L.: Random forests, *Mach. Learn.*, 45, 5–32, 2001.
- Chang, L.-L., Dwivedi, R., Knowles, J. F., Fang, Y.-H., Niu, G.-Y., Pelletier, J. D., Rasmussen, C., Durcik, M., Barron-Gafford, G. A., and Meixner, T.: Why Do Large-Scale Land Surface Models Produce a Low Ratio of Transpiration to Evapotranspiration?, *J. Geophys. Res.-Atmos.*, 123, 9109–9130, 2018.
- Chen, Y., Xia, J., Liang, S., Feng, J., Fisher, J. B., Li, X., Li, X., Liu, S., Ma, Z., and Miyata, A.: Comparison of satellite-based evapotranspiration models over terrestrial ecosystems in China, *Remote Sens. Environ.*, 140, 279–293, 2014.
- Chen, Y., Feng, X., Fu, B., Shi, W., Yin, L., and Lv, Y.: Recent global cropland water consumption constrained by observations, *Water Resour. Res.*, 55, 3708–3738, 2019.
- Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., and Mackay, D. S.: Improving the representation of hydrologic processes in Earth System Models, *Water Resour. Res.*, 51, 5929–5956, 2015.
- Cleugh, H. A., Leuning, R., Mu, Q., and Running, S. W.: Regional evaporation estimates from flux tower and MODIS satellite data, *Remote Sens. Environ.*, 106, 285–304, 2007.
- Dai, A.: Precipitation characteristics in eighteen coupled climate models, *J. Climate*, 19, 4605–4630, 2006.
- Decker, M.: Development and evaluation of a new soil moisture and runoff parameterization for the CABLE LSM including subgrid-scale processes, *J. Adv. Model. Earth Sy.*, 7, 1788–1809, 2015.
- De Kauwe, M. G., Kala, J., Lin, Y.-S., Pitman, A. J., Medlyn, B. E., Duursma, R. A., Abramowitz, G., Wang, Y.-P., and Miralles, D. G.: A test of an optimal stomatal conductance scheme within the CABLE land surface model, *Geosci. Model Dev.*, 8, 431–452, <https://doi.org/10.5194/gmd-8-431-2015>, 2015.
- d’Orgeval, T., Polcher, J., and de Rosnay, P.: Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes, *Hydrol. Earth Syst. Sci.*, 12, 1387–1401, <https://doi.org/10.5194/hess-12-1387-2008>, 2008.
- Dou, X. and Yang, Y.: Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems, *Comput. Electron. Agr.*, 148, 95–106, 2018.
- Douville, H., Ribes, A., Decharme, B., Alkama, R., and Sheffield, J.: Anthropogenic influence on multidecadal changes in reconstructed global evapotranspiration, *Nat. Clim. Change*, 3, 59–62, 2013.
- Famiglietti, J. and Wood, E. F.: Evapotranspiration and runoff from large land areas: Land surface hydrology for atmospheric general circulation models, *Surv. Geophys.*, 12, 179–204, 1991.
- Fatichi, S., Ivanov, V. Y., and Caporali, E.: Investigating Interannual Variability of Precipitation at the Global Scale: Is There a Connection with Seasonality?, *J. Climate*, 25, 5512–5523, 2012.
- Feng, Y., Cui, N., Gong, D., Zhang, Q., and Zhao, L.: Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling, *Agr. Water Manage.*, 193, 163–173, 2017.
- Fisher, J. B., Tu, K. P., and Baldocchi, D. D.: Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites, *Remote Sens. Environ.*, 112, 901–919, 2008.
- Fisher, J. B., Malhi, Y., Bonal, D., Da Rocha, H. R., De Araujo, A. C., Gamo, M., Goulden, M. L., Hirano, T., Huete, A. R., and Kondo, H.: The land–atmosphere water flux in the tropics, *Glob. Change Biol.*, 15, 2694–2714, 2009.
- Fisher, J. B., Melton, F., Middleton, E., Hain, C., Anderson, M., Allen, R., McCabe, M. F., Hook, S., Baldocchi, D., and Townsend, P. A.: The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources, *Water Resour. Res.*, 53, 2618–2626, 2017.
- Fu, B. P.: On the calculation of the evaporation from land surface, *Sci. Atmos. Sin.*, 5, 23–31, 1981.
- Fu, G., Charles, S. P., and Yu, J.: A critical overview of pan evaporation trends over the last 50 years, *Climatic Change*, 97, 193–214, 2009.
- Gedney, N., Cox, P., Betts, R., Boucher, O., Huntingford, C., and Stott, P.: Detection of a direct carbon dioxide effect in continental river runoff records, *Nature*, 439, 835–838, 2006.
- Glenn, E. P., Nagler, P. L., and Huete, A. R.: Vegetation index methods for estimating evapotranspiration by remote sensing, *Surv. Geophys.*, 31, 531–555, 2010.
- Gocic, M., Petković, D., Shamshirband, S., and Kamsin, A.: Comparative analysis of reference evapotranspiration equations mod-

- elling by extreme learning machine, *Comput. Electron. Agr.*, 127, 56–63, 2016.
- Good, S. P., Noone, D., and Bowen, G.: Hydrologic connectivity constrains partitioning of global terrestrial water fluxes, *Science*, 349, 175–177, 2015.
- Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nédélec, S., Ottlé, C., Jornet-Puig, A., Bastos, A., Laurent, P., Goll, D., Bowring, S., Chang, J., Guenet, B., Tifafi, M., Peng, S., Krinner, G., Ducharme, A., Wang, F., Wang, T., Wang, X., Wang, Y., Yin, Z., Lauerwald, R., Joetzjer, E., Qiu, C., Kim, H., and Ciais, P.: ORCHIDEE-MICT (v8.4.1), a land surface model for the high latitudes: model description and validation, *Geosci. Model Dev.*, 11, 121–163, <https://doi.org/10.5194/gmd-11-121-2018>, 2018.
- Haverd, V., Smith, B., Nieradzik, L., Briggs, P. R., Woodgate, W., Trudinger, C. M., Canadell, J. G., and Cuntz, M.: A new version of the CABLE land surface model (Subversion revision r4601) incorporating land use and land cover change, woody vegetation demography, and a novel optimisation-based approach to plant coordination of photosynthesis, *Geosci. Model Dev.*, 11, 2995–3026, <https://doi.org/10.5194/gmd-11-2995-2018>, 2018.
- Heijmans, M. M., Arp, W. J., and Berendse, F.: Effects of elevated CO<sub>2</sub> and vascular plants on evapotranspiration in bog vegetation, *Glob. Change Biol.*, 7, 817–827, 2001.
- Hu, Z., Shi, H., Cheng, K., Wang, Y. P., Piao, S., Li, Y., Zhang, L., Xia, J., Zhou, L., and Yuan, W.: Joint structural and physiological control on the interannual variation in productivity in a temperate grassland: A data-model comparison, *Glob. Change Biol.*, 24, 2965–2979, 2018.
- Huang, S., Bartlett, P., and Arain, M. A.: Assessing nitrogen controls on carbon, water and energy exchanges in major plant functional types across North America using a carbon and nitrogen coupled ecosystem model, *Ecol. Model.*, 323, 12–27, 2016.
- Hulley, G., Hook, S., Fisher, J., and Lee, C.: ECOSTRESS, A NASA Earth-Ventures Instrument for studying links between the water cycle and plant health over the diurnal cycle, 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 23–28 July 2017, Fort Worth, USA, 5494–5496, 2017.
- Ito, A.: Evaluation of the impacts of defoliation by tropical cyclones on a Japanese forest's carbon budget using flux data and a process-based model, *J. Geophys. Res.-Biogeo.*, 115, G04013, <https://doi.org/10.1029/2010JG001314>, 2010.
- Jackson, R., Reginato, R., and Idso, S.: Wheat canopy temperature: a practical tool for evaluating water requirements, *Water Resour. Res.*, 13, 651–656, 1977.
- Jimenez, C., Prigent, C., Mueller, B., Seneviratne, S. I., McCabe, M., Wood, E., Rossow, W., Balsamo, G., Betts, A., and Dirmeyer, P.: Global intercomparison of 12 land surface heat flux estimates, *J. Geophys. Res.-Atmos.*, 116, D02102, <https://doi.org/10.1029/2010JD014545>, 2011.
- Jung, M.: Latent heat flux on land, available at: <https://www.bgc-jena.mpg.de/geodb/projects/FileDetails.php>, last access: 27 March 2020.
- Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, *Biogeosciences*, 6, 2001–2013, <https://doi.org/10.5194/bg-6-2001-2009>, 2009.
- Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., and De Jeu, R.: Recent decline in the global land evapotranspiration trend due to limited moisture supply, *Nature*, 467, 951–954, 2010.
- Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., and Chen, J.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, *J. Geophys. Res.-Biogeo.*, 116, G00J07, <https://doi.org/10.1029/2010JG001566>, 2011.
- Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., Arneth, A., Camps-Valls, G., Ciais, P., and Friedlingstein, P.: Compensatory water effects link yearly global land CO<sub>2</sub> sink changes to temperature, *Nature*, 541, 516–520, 2017.
- Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., and Reichstein, M.: The FLUXCOM ensemble of global land-atmosphere energy fluxes, *Scientific Data*, 6, 74, <https://doi.org/10.1038/s41597-019-0076-8>, 2019.
- Ke, Y., Im, J., Park, S., and Gong, H.: Downscaling of MODIS One kilometer evapotranspiration using Landsat-8 data and machine learning approaches, *Remote Sens.*, 8, 215, <https://doi.org/10.3390/rs8030215>, 2016.
- Keller, K. M., Lienert, S., Bozbiyik, A., Stocker, T. F., Churakova (Sidorova), O. V., Frank, D. C., Klesse, S., Koven, C. D., Leuenberger, M., Riley, W. J., Saurer, M., Siegwolf, R., Weigt, R. B., and Joos, F.: 20th century changes in carbon isotopes and water-use efficiency: tree-ring-based evaluation of the CLM4.5 and LPX-Bern models, *Biogeosciences*, 14, 2641–2673, <https://doi.org/10.5194/bg-14-2641-2017>, 2017.
- Kendall, M. G.: Rank correlation methods, Charles Griffin, London, UK, 196 pp., 1955.
- Khan, M. S., Liaqat, U. W., Baik, J., and Choi, M.: Stand-alone uncertainty characterization of GLEAM, GLDAS and MOD16 evapotranspiration products using an extended triple collocation approach, *Agr. Forest Meteorol.*, 252, 256–268, 2018.
- Kiş, Ö. and Tombul, M.: Modeling monthly pan evaporation using fuzzy genetic approach, *J. Hydrol.*, 477, 203–212, 2013.
- Kisi, O., Sanikhani, H., Zounemat-Kermani, M., and Niazi, F.: Long-term monthly evapotranspiration modeling by several data-driven methods without climatic data, *Comput. Electron. Agr.*, 115, 66–77, 2015.
- Knauer, J., Werner, C., and Zaehle, S.: Evaluating stomatal models and their atmospheric drought response in a land surface scheme: A multibiome analysis, *J. Geophys. Res.-Biogeo.*, 120, 1894–1911, 2015.
- Lawrence, D. M., Thornton, P. E., Oleson, K. W., and Bonan, G. B.: The partitioning of evapotranspiration into transpiration, soil evaporation, and canopy evaporation in a GCM: Impacts on land-atmosphere interaction, *J. Hydrometeorol.*, 8, 862–880, 2007.
- Leipprand, A. and Gerten, D.: Global effects of doubled atmospheric CO<sub>2</sub> content on evapotranspiration, soil moisture and runoff under potential natural vegetation, *Hydrolog. Sci. J.*, 51, 171–185, 2006.
- Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell,

- J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., Bakker, D. C. E., Barbero, L., Becker, M., Betts, R. A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Cosca, C. E., Cross, J., Currie, K., Gasser, T., Harris, I., Hauck, J., Haverd, V., Houghton, R. A., Hunt, C. W., Hurtt, G., Ilyina, T., Jain, A. K., Kato, E., Kautz, M., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lima, I., Lombardozzi, D., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., Nojiri, Y., Padin, X. A., Peregon, A., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Reimer, J., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Viovy, N., Vuichard, N., Walker, A. P., Watson, A. J., Wiltshire, A. J., Zaehle, S., and Zhu, D.: Global Carbon Budget 2017, *Earth Syst. Sci. Data*, 10, 405–448, <https://doi.org/10.5194/essd-10-405-2018>, 2018.
- Leuning, R., Zhang, Y., Rajaud, A., Cleugh, H., and Tu, K.: A simple surface conductance model to estimate regional evaporation using MODIS leaf area index and the Penman-Monteith equation, *Water Resour. Res.*, 44, W10419, <https://doi.org/10.1029/2007WR006562>, 2008.
- Li, J., Wang, Y. P., Duan, Q., Lu, X., Pak, B., Wiltshire, A., Robertson, E., and Ziehn, T.: Quantification and attribution of errors in the simulated annual gross primary production and latent heat fluxes by two global land surface models, *J. Adv. Model. Earth Sy.*, 8, 1270–1288, 2016.
- Li, L., Wang, Y. P., Yu, Q., Pak, B., Eamus, D., Yan, J., Gorsel, E., and Baker, I. T.: Improving the responses of the Australian community land surface model (CABLE) to seasonal drought, *J. Geophys. Res.-Biogeo.*, 117, G04002, <https://doi.org/10.1029/2012JG002038>, 2012.
- Li, Z.-L., Tang, R., Wan, Z., Bi, Y., Zhou, C., Tang, B., Yan, G., and Zhang, X.: A review of current methodologies for regional evapotranspiration estimation from remotely sensed data, *Sensors*, 9, 3801–3853, 2009.
- Lian, X., Piao, S., Huntingford, C., Li, Y., Zeng, Z., Wang, X., Ciais, P., McVicar, T. R., Peng, S., and Ottlé, C.: Partitioning global land evapotranspiration using CMIP5 models constrained by observations, *Nat. Clim. Change*, 8, 640–646, 2018.
- Liu, M., Tian, H., Chen, G., Ren, W., Zhang, C., and Liu, J.: Effects of Land-Use and Land-Cover Change on Evapotranspiration and Water Yield in China During 1900–2000, *J. Am. Water Resour. As.*, 44, 1193–1207, 2008.
- Ma, Y., Liu, S., Song, L., Xu, Z., Liu, Y., Xu, T., and Zhu, Z.: Estimation of daily evapotranspiration and irrigation water efficiency at a Landsat-like scale for an arid irrigation area using multi-source remote sensing data, *Remote Sens. Environ.*, 216, 715–734, 2018.
- Mann, H. B.: Nonparametric tests against trend, *Econometrica*, 1945, 245–259, 1945.
- Mao, J., Fu, W., Shi, X., Ricciuto, D. M., Fisher, J. B., Dickinson, R. E., Wei, Y., Shem, W., Piao, S., and Wang, K.: Disentangling climatic and anthropogenic controls on global terrestrial evapotranspiration trends, *Environ. Res. Lett.*, 10, 094008, <https://doi.org/10.1088/1748-9326/10/9/094008>, 2015.
- Mao, J., Ribes, A., Yan, B., Shi, X., Thornton, P. E., Séférian, R., Ciais, P., Myneni, R. B., Douville, H., and Piao, S.: Human-induced greening of the northern extratropical land surface, *Nat. Clim. Change*, 2016, 959–963, 2016.
- Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, *Geosci. Model Dev.*, 10, 1903–1925, <https://doi.org/10.5194/gmd-10-1903-2017>, 2017.
- Maselli, F., Papale, D., Chiesi, M., Matteucci, G., Angeli, L., Raschi, A., and Seufert, G.: Operational monitoring of daily evapotranspiration by the combination of MODIS NDVI and ground meteorological data: Application and evaluation in Central Italy, *Remote Sens. Environ.*, 152, 279–290, 2014.
- Medlyn, B. E., Zaehle, S., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hanson, P. J., Hickler, T., Jain, A. K., Luo, Y., and Parton, W.: Using ecosystem experiments to improve vegetation models, *Nat. Clim. Change*, 5, 528–534, 2015.
- Melton, J. R. and Arora, V. K.: Competition between plant functional types in the Canadian Terrestrial Ecosystem Model (CTEM) v. 2.0, *Geosci. Model Dev.*, 9, 323–361, <https://doi.org/10.5194/gmd-9-323-2016>, 2016.
- Miralles, D. G.: GLEAM v3, available at: <https://www.gleam.eu/>, last access: 27 March 2020.
- Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, *Hydrol. Earth Syst. Sci.*, 15, 453–469, <https://doi.org/10.5194/hess-15-453-2011>, 2011.
- Miralles, D. G., van den Berg, M., Gash, J., Parinussa, R., Jeu, R., Beck, H., Holmes, T., Jimenez, C., Verhoest, N., and Dorigo, W.: El Niño-La Niña cycle and recent trends in continental evaporation, *Nat. Clim. Change*, 4, 122–126, 2014.
- Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., McCabe, M. F., Hirschi, M., Martens, B., Dolman, A. J., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D.: The WACMOS-ET project – Part 2: Evaluation of global terrestrial evaporation data sets, *Hydrol. Earth Syst. Sci.*, 20, 823–842, <https://doi.org/10.5194/hess-20-823-2016>, 2016.
- Monteith, J. L.: Evaporation and environment, *Symposia of the society for experimental biology*, Cambridge University Press (CUP), Cambridge, UK, 19, 205–234, 1965.
- Monteith, J. L.: Accommodation between transpiring vegetation and the convective boundary layer, *J. Hydrol.*, 166, 251–263, 1995.
- Mu, Q., Heinsch, F. A., Zhao, M., and Running, S. W.: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, *Remote Sens. Environ.*, 111, 519–536, 2007.
- Mu, Q., Zhao, M., and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sens. Environ.*, 115, 1781–1800, 2011.
- Mueller, B., Seneviratne, S. I., Jimenez, C., Corti, T., Hirschi, M., Balsamo, G., Ciais, P., Dirmeyer, P., Fisher, J., and Guo, Z.: Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations, *Geophys. Res. Lett.*, 38, L06402, <https://doi.org/10.1029/2010GL046230>, 2011.
- Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung, M., Ludwig, F., Maignan, F., Miralles, D. G., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K., Wood, E. F., Zhang, Y., and Seneviratne, S.

- I.: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis, *Hydrol. Earth Syst. Sci.*, 17, 3707–3720, <https://doi.org/10.5194/hess-17-3707-2013>, 2013.
- Murray-Tortarolo, G., Anav, A., Friedlingstein, P., Sitch, S., Piao, S., Zhu, Z., Poulter, B., Zaehle, S., Ahlström, A., and Lomas, M.: Evaluation of land surface models in reproducing satellite-derived LAI over the high-latitude northern hemisphere. Part I: Uncoupled DGVMs, *Remote Sens.*, 5, 4819–4838, 2013.
- Nagler, P. L., Scott, R. L., Westenburg, C., Cleverly, J. R., Glenn, E. P., and Huete, A. R.: Evapotranspiration on western US rivers estimated using the Enhanced Vegetation Index from MODIS and data from eddy covariance and Bowen ratio flux towers, *Remote Sens. Environ.*, 97, 337–351, 2005.
- Nepstad, D., Lefebvre, P., Lopes da Silva, U., Tomasella, J., Schlesinger, P., Solórzano, L., Moutinho, P., Ray, D., and Guerreira Benito, J.: Amazon drought and its implications for forest flammability and tree growth: A basin-wide analysis, *Glob. Change Biol.*, 10, 704–717, 2004.
- Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., and Rosero, E.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.-Atmos.*, 116, D12109, <https://doi.org/10.1029/2010JD015139>, 2011.
- Noilhan, J. and Mahfouf, J.-F.: The ISBA land surface parameterisation scheme, *Global Planet. Change*, 13, 145–159, 1996.
- Norman, J. M. and Becker, F.: Terminology in thermal infrared remote sensing of natural surfaces, *Remote Sens. Rev.*, 12, 159–173, 1995.
- Oki, T. and Kanae, S.: Global hydrological cycles and world water resources, *Science*, 313, 1068–1072, 2006.
- Oleson, K. W., Lawrence, D. M., Gordon, B., Flanner, M. G., Kluzek, E., Peter, J., Levis, S., Swenson, S. C., Thornton, E., and Feddes, J.: Technical description of version 4.0 of the Community Land Model (CLM), NCAR technical note, National Center for Atmospheric Research, Boulder, CO, USA, 257 pp., 2010.
- Pagán, R. B., Maes, H. W., Gentile, P., Martens, B., and Miralles, G. D.: Exploring the Potential of Satellite Solar-Induced Fluorescence to Constrain Global Transpiration Estimates, *Remote Sens.*, 11, 413, <https://doi.org/10.3390/rs11040413>, 2019.
- Pan, N., Feng, X., Fu, B., Wang, S., Ji, F., and Pan, S.: Increasing global vegetation browning hidden in overall vegetation greening: Insights from time-varying trends, *Remote Sens. Environ.*, 214, 59–72, 2018.
- Pan, S., Tian, H., Dangal, S. R., Yang, Q., Yang, J., Lu, C., Tao, B., Ren, W., and Ouyang, Z.: Responses of global terrestrial evapotranspiration to climate change and increasing atmospheric CO<sub>2</sub> in the 21st century, *Earth's Future*, 3, 15–35, 2015.
- Pan, S., Chen, G., Ren, W., Dangal, S. R. S., Banger, K., Yang, J., Tao, B., and Tian, H.: Responses of global terrestrial water use efficiency to climate change and rising atmospheric CO<sub>2</sub> concentration in the twenty-first century, *Int. J. Digit. Earth*, 11, 558–582, 2018.
- Peng, S., Piao, S., Ciais, P., Myneni, R. B., Chen, A., Chevalier, F., Dolman, A. J., Janssens, I. A., Penuelas, J., and Zhang, G.: Asymmetric effects of daytime and night-time warming on Northern Hemisphere vegetation, *Nature*, 501, 88–92, 2013.
- Peterson, T. C., Golubev, V. S., and Groisman, P. Y.: Evaporation losing its strength, *Nature*, 377, 687–688, 1995.
- Piao, S., Ciais, P., Huang, Y., Shen, Z., Peng, S., Li, J., Zhou, L., Liu, H., Ma, Y., and Ding, Y.: The impacts of climate change on water resources and agriculture in China, *Nature*, 467, 43–51, 2010.
- Pinzon, J. E. and Tucker, C. J.: GIMMS NDVI3gV1, available at: <https://ecocast.arc.nasa.gov/data/pub/gimms/3g.v1/>, last access: 27 March 2020.
- Poon, P. and Kinoshita, A.: Estimating Evapotranspiration in a Post-Fire Environment Using Remote Sensing and Machine Learning, *Remote Sens.*, 10, 1728, <https://doi.org/10.3390/rs10111728>, 2018.
- Priestley, C. and Taylor, R.: On the assessment of surface heat flux and evaporation using large-scale parameters, *Mon. Weather Rev.*, 100, 81–92, 1972.
- Purdy, A. J., Fisher, J. B., Goulden, M. L., Colliander, A., Halverson, G., Tu, K., and Famiglietti, J. S.: SMAP soil moisture improves global evapotranspiration, *Remote Sens. Environ.*, 219, 1–14, 2018.
- Reichstein, M., Besnard, S., Carvalhais, N., Gans, F., Jung, M., Kraft, B., and Mahecha, M.: Modelling Landsurface Time-Series with Recurrent Neural Nets, *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, 22–27 July 2018, Valencia, Spain, 7640–7643, 2018.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, *Nature*, 566, 195–204, 2019.
- Restrepo-Coupe, N., Levine, N. M., Christoffersen, B. O., Albert, L. P., Wu, J., Costa, M. H., Galbraith, D., Imbuzeiro, H., Martins, G., and da Araujo, A. C.: Do dynamic global vegetation models capture the seasonality of carbon fluxes in the Amazon basin? A data-model intercomparison, *Glob. Change Biol.*, 23, 191–208, 2017.
- Roderick, M. L. and Farquhar, G. D.: The cause of decreased pan evaporation over the past 50 years, *Science*, 298, 1410–1411, 2002.
- Rouholahnejad, E. and Martens, B.: Improving global estimates of terrestrial evaporation by integrating remotely-sensed observations of changes in terrestrial water storage (GRACE) into a global evaporation model (GLEAM), *AGU Fall Meeting*, 10–14 December 2018, Washington, D.C., USA, 2018.
- Running, S. W.: A measurable planetary boundary for the biosphere, *Science*, 337, 1458–1459, 2012.
- Running, S. W.: MOD16A\_MONTHLY.MERRA\_GMAO\_1kmALB, available at: [http://files.nts.gov.umt.edu/data/NTSG\\_Products/MOD16/](http://files.nts.gov.umt.edu/data/NTSG_Products/MOD16/), last access: 27 March 2020.
- Sayemuzzaman, M. and Jha, M. K.: Seasonal and annual precipitation time series trend analysis in North Carolina, United States, *Atmos. Res.*, 137, 183–194, 2014.
- Semmens, K. A., Anderson, M. C., Kustas, W. P., Gao, F., Alfieri, J. G., McKee, L., Prueger, J. H., Hain, C. R., Cammalleri, C., Yang, Y., Xia, T., Sanchez, L., Mar Alsina, M., and Vélez, M.: Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach, *Remote Sens. Environ.*, 185, 155–170, 2016.
- Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, *J. American Stat. Assoc.*, 63, 1379–1389, 1968.
- Shrestha, N. and Shukla, S.: Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a

- sub-tropical environment, *Agr. Forest Meteorol.*, 200, 172–184, 2015.
- Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., and Portmann, F. T.: Groundwater use for irrigation – a global inventory, *Hydrol. Earth Syst. Sci.*, 14, 1863–1880, <https://doi.org/10.5194/hess-14-1863-2010>, 2010.
- Simmons, A., Willett, K., Jones, P., Thorne, P., and Dee, D.: Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets, *J. Geophys. Res.-Atmos.*, 115, D01110, <https://doi.org/10.1029/2009JD012442>, 2010.
- Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J., Levis, S., Lucht, W., and Sykes, M. T.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, *Glob. Change Biol.*, 9, 161–185, 2003.
- Smith, B.: LPJ-GUESS-an ecosystem modelling framework, Department of Physical Geography and Ecosystems Analysis, INES, Sölvegatan, 12, 22362, 2001.
- Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., de Vries, W., de Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B., and Sörlin, S.: Planetary boundaries: Guiding human development on a changing planet, *Science*, 347, 1259855, <https://doi.org/10.1126/science.1259855>, 2015.
- Stoy, P. C., El-Madany, T. S., Fisher, J. B., Gentile, P., Gerken, T., Good, S. P., Klosterhalfen, A., Liu, S., Miralles, D. G., Perez-Priego, O., Rigden, A. J., Skaggs, T. H., Wohlfahrt, G., Anderson, R. G., Coenders-Gerrits, A. M. J., Jung, M., Maes, W. H., Mammarella, I., Mauder, M., Migliavacca, M., Nelson, J. A., Poyatos, R., Reichstein, M., Scott, R. L., and Wolf, S.: Reviews and syntheses: Turning the challenges of partitioning ecosystem evaporation and transpiration into opportunities, *Biogeosciences*, 16, 3747–3775, <https://doi.org/10.5194/bg-16-3747-2019>, 2019.
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K. L.: A review of global precipitation data sets: Data sources, estimation, and intercomparisons, *Rev. Geophys.*, 56, 79–107, 2018.
- Sun, Y., Frankenberg, C., Wood, J. D., Schimel, D. S., Jung, M., Guanter, L., Drewry, D. T., Verma, M., Porcar-Castell, A., Griffiths, T. J., Gu, L., Magney, T. S., Köhler, P., Evans, B., and Yuen, K.: OCO-2 advances photosynthesis observation from space via solar-induced chlorophyll fluorescence, *Science*, 358, eaam5747, <https://doi.org/10.1126/science.aam5747>, 2017.
- Swann, A. L., Hoffman, F. M., Koven, C. D., and Randerson, J. T.: Plant responses to increasing CO<sub>2</sub> reduce estimates of climate impacts on drought severity, *P. Natl. Acad. Sci. USA*, 113, 10019–10024, 2016.
- Szutu, D. J. and Papuga, S. A.: Year-Round Transpiration Dynamics Linked With Deep Soil Moisture in a Warm Desert Shrubland, *Water Resour. Res.*, 55, 5679–5695, 2019.
- Tabari, H., Martinez, C., Ezani, A., and Talaei, P. H.: Applicability of support vector machines and adaptive neurofuzzy inference system for modeling potato crop evapotranspiration, *Irrigation Sci.*, 31, 575–588, 2013.
- Talsma, C. J., Good, S., Miralles, D., Fisher, J., Martens, B., Jimenez, C., and Purdy, A.: Sensitivity of Evapotranspiration Components in Remote Sensing-Based Models, *Remote Sens.*, 10, 1601, <https://doi.org/10.3390/rs10101601>, 2018a.
- Talsma, C. J., Good, S. P., Jimenez, C., Martens, B., Fisher, J. B., Miralles, D. G., McCabe, M. F., and Purdy, A. J.: Partitioning of evapotranspiration in remote sensing-based models, *Agr. Forest Meteorol.*, 260–261, 131–143, 2018b.
- Tian, H., Chen, G., Liu, M., Zhang, C., Sun, G., Lu, C., Xu, X., Ren, W., Pan, S., and Chappelka, A.: Model estimates of net primary productivity, evapotranspiration, and water use efficiency in the terrestrial ecosystems of the southern United States during 1895–2007, *Forest Ecol. Manage.*, 259, 1311–1327, 2010.
- Trenberth, K. E., Fasullo, J. T., and Kiehl, J.: Earth's global energy budget, *B. Am. Meteorol. Soc.*, 90, 311–324, 2009.
- Van Dijk, A. and Bruijnzeel, L. A.: Modelling rainfall interception by vegetation of variable density using an adapted analytical model. Part I. Model description, *J. Hydrol.*, 247, 230–238, 2001.
- Vinukollu, R. K., Meynadier, R., Sheffield, J., and Wood, E. F.: Multi-model, multi-sensor estimates of global evapotranspiration: Climatology, uncertainties and trends, *Hydrol. Process.*, 25, 3993–4010, 2011a.
- Vinukollu, R. K., Wood, E. F., Ferguson, C. R., and Fisher, J. B.: Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches, *Remote Sens. Environ.*, 115, 801–823, 2011b.
- Wang, K. and Dickinson, R. E.: A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, *Rev. Geophys.*, 50, RG2005, <https://doi.org/10.1029/2011RG000373>, 2012.
- Wang, K., Dickinson, R. E., Wild, M., and Liang, S.: Evidence for decadal variation in global terrestrial evapotranspiration between 1982 and 2002: 1. Model development, *J. Geophys. Res.-Atmos.*, 115, D20112, <https://doi.org/10.1029/2009JD013671>, 2010.
- Wartenburger, R., Seneviratne, S. I., Hirschi, M., Chang, J., Ciais, P., Deryng, D., Elliott, J., Folberth, C., Gosling, S. N., and Gudmundsson, L.: Evapotranspiration simulations in ISIMIP2a-Evaluation of spatio-temporal characteristics with a comprehensive ensemble of independent datasets, *Environ. Res. Lett.*, 13, 075001, <https://doi.org/10.1088/1748-9326/aac4bb>, 2018.
- Willett, K. M., Gillett, N. P., Jones, P. D., and Thorne, P. W.: Attribution of observed surface humidity changes to human influence, *Nature*, 449, 710–712, 2007.
- Xu, D., Agee, E., Wang, J., and Ivanov, V. Y.: Estimation of Evapotranspiration of Amazon Rainforest Using the Maximum Entropy Production Method, *Geophys. Res. Lett.*, 46, 1402–1412, 2019.
- Xu, T., Guo, Z., Liu, S., He, X., Meng, Y., Xu, Z., Xia, Y., Xiao, J., Zhang, Y., Ma, Y., and Song, L.: Evaluating Different Machine Learning Methods for Upscaling Evapotranspiration from Flux Towers to the Regional Scale, *J. Geophys. Res.-Atmos.*, 123, 8674–8690, 2018.
- Yao, Y., Liang, S., Cheng, J., Liu, S., Fisher, J. B., Zhang, X., Jia, K., Zhao, X., Qin, Q., and Zhao, B.: MODIS-driven estimation of terrestrial latent heat flux in China based on a modified Priestley–Taylor algorithm, *Agr. Forest Meteorol.*, 171, 187–202, 2013.
- Yao, Y., Liang, S., Li, X., Liu, S., Chen, J., Zhang, X., Jia, K., Jiang, B., Xie, X., and Munier, S.: Assessment and simulation of global terrestrial latent heat flux by synthesis of CMIP5 climate models



- and surface eddy covariance observations, *Agr. Forest Meteorol.*, 223, 151–167, 2016.
- Yao, Y., Liang, S., Li, X., Chen, J., Liu, S., Jia, K., Zhang, X., Xiao, Z., Fisher, J. B., and Mu, Q.: Improving global terrestrial evapotranspiration estimation using support vector machine by integrating three process-based algorithms, *Agr. Forest Meteorol.*, 242, 55–74, 2017.
- Yuan, W., Liu, S., Liu, H., Randerson, J. T., Yu, G., and Tieszen, L. L.: Impacts of precipitation seasonality and ecosystem types on evapotranspiration in the Yukon River Basin, Alaska, *Water Resour. Res.*, 46, W02514, <https://doi.org/10.1029/2009WR008119>, 2010.
- Yue, S., Pilon, P., and Cavadias, G.: Power of the Mann–Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series, *J. Hydrol.*, 259, 254–271, 2002.
- Zaehle, S. and Friend, A.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates, *Global Biogeochem. Cy.*, 24, GB1005, <https://doi.org/10.1029/2009GB003521>, 2010.
- Zeng, Z., Piao, S., Lin, X., Yin, G., Peng, S., Ciais, P., and Myneni, R. B.: Global evapotranspiration over the past three decades: estimation based on the water balance equation combined with empirical models, *Environ. Res. Lett.*, 7, 014026, <https://doi.org/10.1088/1748-9326/7/1/014026>, 2012.
- Zeng, Z., Wang, T., Zhou, F., Ciais, P., Mao, J., Shi, X., and Piao, S.: A worldwide analysis of spatiotemporal changes in water balance-based evapotranspiration from 1982 to 2009, *J. Geophys. Res.-Atmos.*, 119, 1186–1202, 2014.
- Zeng, Z., Zhu, Z., Lian, X., Li, L. Z., Chen, A., He, X., and Piao, S.: Responses of land evapotranspiration to Earth's greening in CMIP5 Earth System Models, *Environ. Res. Lett.*, 11, 104006, <https://doi.org/10.1088/1748-9326/11/10/104006>, 2016.
- Zeng, Z., Peng, L., and Piao, S.: Response of terrestrial evapotranspiration to Earth's greening, *Curr. Opin. Env. Sust.*, 33, 9–25, 2018a.
- Zeng, Z., Piao, S., Li, L. Z., Wang, T., Ciais, P., Lian, X., Yang, Y., Mao, J., Shi, X., and Myneni, R. B.: Impact of Earth greening on the terrestrial water cycle, *J. Climate*, 31, 2633–2650, 2018b.
- Zhang, K.: P-LSH ET, available at: [http://files.ntsg.umd.edu/data/ET\\_global\\_monthly/Global\\_8kmResolution/](http://files.ntsg.umd.edu/data/ET_global_monthly/Global_8kmResolution/), last access: 27 March 2020.
- Zhang, K., Kimball, J. S., Nemani, R. R., and Running, S. W.: A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006, *Water Resour. Res.*, 46, W09522, <https://doi.org/10.1029/2009WR008800>, 2010.
- Zhang, K., Kimball, J. S., Nemani, R. R., Running, S. W., Hong, Y., Gourley, J. J., and Yu, Z.: Vegetation greening and climate change promote multidecadal rises of global land evapotranspiration, *Sci. Rep.*, 5, 15956, <https://doi.org/10.1038/srep15956>, 2015.
- Zhang, K., Kimball, J. S., and Running, S. W.: A review of remote sensing based actual evapotranspiration estimation, *WIREs: Water*, 3, 834–853, 2016.
- Zhang, L., Lei, H., Shen, H., Cong, Z., Yang, D., and Liu, T.: Evaluating the Representation of Vegetation Phenology in the Community Land Model 4.5 in a Temperate Grassland, *J. Geophys. Res.-Biogeo.*, 124, 187–210, 2019.
- Zhang, Y.: Monthly global observation-driven Penman-Monteith-Leuning (PML) evapotranspiration and components, available at: <https://data.csiro.au/dap/landingpage?pid=csiro:17375>, last access: 27 March 2020.
- Zhang, Y., Peña-Arancibia, J. L., McVicar, T. R., Chiew, F. H., Vaze, J., Liu, C., Lu, X., Zheng, H., Wang, Y., and Liu, Y. Y.: Multi-decadal trends in global terrestrial evapotranspiration and its components, *Sci. Rep.*, 6, 19124, <https://doi.org/10.1038/srep19124>, 2016.
- Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. R., and Myneni, R. B.: Global data sets of vegetation leaf area index (LAI) 3g and Fraction of Photosynthetically Active Radiation (FPAR) 3g derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the period 1981 to 2011, *Remote Sens.*, 5, 927–948, 2013.
- Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P., and Arneth, A.: Greening of the Earth and its drivers, *Nat. Clim. Change*, 2016, 791–795, 2016.